الجمهورية الجزائرية الدّيمقراطية الشّعبية

وزارة التّعليم العالي والبحث العلمي

جامعة ابن خلدون -تيارت –

كلّية: العلوم الاقتصادية التّجارية وعلوم التّسيير

قسم: العلوم التجارية

دروس محاضرات في مقياس -إحصاء 3-

(التوزيعات الإحصائية)

من إعداد:

- شـــــداد محمد

السنة الجامعية 2024-2025

فهرس المحتويات

1. مقدمــــــة:

نهدف من خلال هذه المطبوعة تقديم دروس محاضرات في مقياس الإحصاء 03 الذي يدرس في طور الليسانس خلال السداسي الثالث لجميع التخصصات. يمكن أن نلخص أهداف هذا المقياس في العناصر التالية:

- التمهيد التطبيقي للنماذج الاقتصادية النظرية وإعطائها صيغة رياضية .
- التعرف على أهم قوانين التوزيعات الاحتمالية المتقطعة والمستمرة
- إكساب الطالب القدرة على تطبيق التوزيعات الاحتمالية لمعالجة وحل المشكلات الاقتصادية والإدارية والاجتماعية.
- استيعاب المتغيرات العشوائية الثنائية المنفصلة والمتصلة وأهم خواصها.
- التعرف على بعض التوزيعات ذات المتغيرين.

بالنظر إلى تفاصيل محتوى مقياس إحصاء 3 نرى أنه ما هو إلا تسمية أخرى **لنظرية التوزيعات الاحتمالية الإحصائية** أين يتعين على الطالب الإلمام بنظريتي الإحصاء والاحتمالات بجميع عناصرهما والمعبر عنهما بمقياسي إحصاء 1 وإحصاء 02. إضافة إلى ذلك، سيكون من الأهمية على الطالب التحكم في الأدوات الرياضية بصفة عامة وبعض المفاهيم الخاصة الأخرى، كالتكاملات، الاشتقاقات، بعض الجبر الخطي...، حتى يكتمل فهمه لعناصر النظرية أو المقياس.

على هذا الأساس ومن خلال طرحنا التالي سوف نحاول التطرق إلى العناصر التالية:

- أهم قوانين التوزيعات الاحتمالية المتقطعة،
- أهم قوانين التوزيعات الاحتمالية المتصلة،
- تقارب بعض التوزيعات الاحتمالية،
- المتغيرات العشوائية الثنائية.

نشير أننا قد تناولنا هذه العناصر مع بعض التفصيل، كبرهنة أهم العلاقات وأهم النظريات، حتى يكتمل الفهم بالنسبة للطالب ويتشكل لديه قاعدة أساسية متينة لفهم المقاييس التي تعتمد على عناصر مقياس إحصاء 3 على غرار مقاييس الاقتصاد القياسي، الاقتصاد القياسي المعمق وتحليل السلاسل الزمنية. كما قمنا بتدعيم كل عنصر بمجموعة من الأمثلة والتمارين. كما لا يفوتنا أن ننوه إلى أن التوزيعين، فيشر *Fisher* وستودنت *Student*، فضلنا أن نفصلهما ضمن محاضرات مقياس إحصاء 4 (نظرية المعاينة). كما لا يفوتني أن أشير إلى أن الهيكل الأساسي لهذه المطبوعة هو مستوحى من الكتاب القيم "Statistical Inference" لمؤلفيه "George Casella" و "Roger L. Berger".

قبل البدء نشير إلى أن التوزيعات الإحصائية تستعمل لنمذجة المجتمعات الإحصائية المعبر عنها بمتغير أو بمجموعة من المتغيرات الإحصائية. نتعامل غالبا مع مجموعة أو عائلة من التوزيعات بدل مع توزيع واحد. هذه العائلة من التوزيعات تكون مؤشرة بمعلمة أو مجموعة من المعالم الأمر الذي يعطينا أفضلية التعامل مع مجموعة من الخصائص للتوزيع بالبقاء مع صيغة رياضية واحدة أو شكل دالي واحد. مثلا، يمكننا القول بأن التوزيع الطبيعي هو الاختيار الأكثر عقلانية ومثالية لنمذجة مجتمع معين، لكن لا يمكننا بدقة تحديد متوسطه μ، أين يكون غير مشخص أو غير محدد، $-\infty < \mu < +\infty$.

سوف نحاول من خلال ما يلي عرض أهم التوزيعات الإحصائية، المتقطعة والمستمرة، ، أحادية وثنائية المتغيرات، بحيث نقدم من أجل كل توزيع متوسطه وتباينه والعديد من قياساته المهمة والمفيدة لفهمه، كما سنشير أيضا إلى أهم تطبيقاها الخاصة والعلاقات التي تربط بينهم.

2. التوزيعات الاحتمالية المتقطعة:

نقول عن متغير عشوائي X بأن له (أو يتبع) توزيع متقطع إذاكانت مجموعة تعريفه (فضاء تعريفه) قابلة للعد حيث تأخذ قيم صحيحة.

2.أ. توزيع أحادي الشكل المتقطع ***Discrete Uniform Distribution***:

نقول عن متغير عشوائي X أن لديه أو يتبع توزيع أحادي الشكل، ونكتب $X \sim U(1,N)$، إذاكان:

$$P(X = x/N) = \frac{1}{N}, x = 1,2, \dots, N \dots \dots (2.1)$$

أين N هو عدد صحيح موجب. خاصية هذا التوزيع الأساسية أنه يعطي قيمة كثافة احتمالية (احتمال في الحالة المتقطعة) متساوية لكل قيم مجموعة التعريف أو المخرجات $1,2, \dots, N$.

ملاحظة حول الترميز:

عندما نتعامل مع التوزيعات المعلمية، مثلما هو الحال في أغلب الحالات، التوزيع هو دالة تابعة لقيم المعالم. من أجل إعطاء أهمية لهذا الأمر وتتبع أثر المعالم، سوف نكتب صيغ دوال كثافتها الاحتمالية متبوعة بالرمز "/" (بحيث), هذا الترميز سيستعمل أيضا في صيغ دوال الكثافة التراكمية وفي حالات أخرى لما يستدعي الأمر ذلك. عندما لا يكون هناك مكان للبس، يمكن إهمال هذا الأمر خصوصا لتفادي إثقال العلاقات بالرموز.

لحساب متوسط وتباين المتغير X، نذكر بالعلاقات:

$$\sum_{i=1}^{K} i^2 = \frac{k(k+1)(2k+1)}{6} \quad \text{و} \quad \sum_{i=1}^{K} i = \frac{k(k+1)}{2}$$

يصبح لدينا إذا:

$$E(X) = \sum_{x=1}^{N} x\, P(X = x/N) = \sum_{x=1}^{N} x \frac{1}{N} = \frac{N+1}{2}$$

و

$$E(X^2) = \sum_{x=1}^{N} x^2\, P(X = x/N) = \sum_{x=1}^{N} x^2 \frac{1}{N} = \frac{(N+1)(2N+1)}{6}$$

ومنه

$$V(X) = E(X^2) - E(X)^2 = \frac{(N+1)(2N+1)}{6} - \left(\frac{N+1}{2}\right)^2$$

$$= \frac{(N+1)(N-1)}{12}$$

يمكن تعميم هذا التوزيع على أي مجال تعريف من الشكل $N_0, N_0+1, \dots, N_1$ وتكتب دالته الاحتمالية على النحو التالي:

$$P(X = x / N_0\,, N_1) = \frac{1}{(N_1 - N_0 + 1)}.$$

2.ب. التوزيع فوق الهندسي *Hypergeometric Distribution*:

التوزيع فوق الهندسي له عدة تطبيقات في المجتمعات المنتهية وأحسن طريقة لفهمه هي من خلال المثال التقليدي لنموذج الكيس.

ليكن لدينا كيس به N كرة متماثلة، منها M حمراء و $N-M$ خضراء. نسحب منه عشوائيا K كرة دفعة واحدة. ما هو احتمال أن نسحب بالضبط x كرة حمراء؟

عدد العينات ذات الحجم K الممكن تشكيلها من N كرة هو C_N^K، يتطلب أن يكون بها x كرة حمراء، وهذا يمكن تحقيقه ب C_M^x طريقة مختلفة مبقيا C_{N-M}^{K-x} طريقة لرفع حجم العينة إلى K ب $K-x$ كرة خضراء. بالرمز ب X لعدد الكريات البيضاء في العينة ذات الحجم K، التوزيع الفوق هندسي (دالته الاحتمالية) ل X يعطى بالعلاقة:

$$P(X = x / N\,, M, K) = \frac{C_M^x C_{N-M}^{K-x}}{C_N^K}, x = 0,1,2, \dots, K \quad \dots \dots (2.2)$$

ونكتب $X \sim H.G(N, M, K)$. نشير هنا أنه يوجد، ضمنيا من خلال أساس التجربة المتمثل في سحب العينة دفعة واحدة ومن خلال العلاقة (2.2) قيود إضافية على مجموعة تعريف X. هذه القيود يعبر عنها بتعريف معاملات ثنائي الحد C_n^r التي تكون معرفة فقط لما يكون $n \geq r$. إذا، بالإضافة، مجموعة تعريف X هي مقيدة بالمتراجحتين :

$$M \geq x \text{ و } N - M \geq K - x$$

الممكن جمعهما بالعلاقة:

$$M - (N - K) \leq x \leq M$$

في أغلب الحالات العملية يكون K أصغر مقارنة ب M و $N-M$، بحيث يكون المجال $0 \leq x \leq K$ محتوى في المجال القيدي السابق ويكون بالضرورة ملائم لمجموعة تعريف X. صيغة الدالة الاحتمالية للتوزيع فوق الهندسي يكون عموما من الصعب التعامل معها حيث ليس من السهل التحقق من:

$$\sum_{x=0}^{K} P(X = x) = \sum_{x=0}^{k} \frac{C_M^x C_{N-M}^{K-x}}{C_N^K} = 1$$

في الواقع يعطي التوزيع فوق الهندسي مثالا عن حقيقة صعوبة التعامل مع المجتمعات المنتهية (N منتهية).

المتوسط أو الأمل الرياضي للتوزيع الفوق الهندسي يعطى ب:

$$E(X)=\sum_{x=0}^{k}x\frac{C_M^xC_{N-M}^{K-x}}{C_N^K}=\sum_{x=1}^{k}x\frac{C_M^xC_{N-M}^{K-x}}{C_N^K}$$

لحساب هذه العبارة نستعين بالعلاقات:

$$xC_M^x=MC_{M-1}^{x-1}$$

$$C_N^K=\frac{N}{K}C_{N-1}^{K-1}$$

ونتحصل على:

$$E(X)=\sum_{x=0}^{k}x\frac{MC_{M-1}^{x-1}C_{N-M}^{K-x}}{\frac{N}{K}C_{N-1}^{K-1}}=\frac{KM}{N}\sum_{x=1}^{k}\frac{C_{M-1}^{x-1}C_{N-M}^{K-x}}{C_{N-1}^{K-1}}$$

يمكننا التعرف على الطرف الأيمن من هذه العلاقة على أنه يمثل مجموع احتمالات متغير عشوائي ذو توزيع فوق هندسي آخر معالمه هي $N-1$ ،$M-1$ و $K-1$. يمكن ملاحظة هذا بوضوح بوضع $y=x-1$ وكتابة:

$$\sum_{x=1}^{k}\frac{C_{M-1}^{x-1}C_{N-M}^{K-x}}{C_{N-1}^{K-1}}=\sum_{y=0}^{k-1}\frac{C_{M-1}^{y}C_{(N-1)-(M-1)}^{K-1-y}}{C_{N-1}^{K-1}}$$

$$=\sum_{y=0}^{K-1}P(Y=y/N-1,M-1,K-1)=1$$

أين Y كما أشرنا هو متغير عشوائي يخضع للتوزيع الفوق الهندسي ذو معالم $N-1,M-1,K-1$ ، أي $Y\backsim H.G(N-1,M-1,K-1)$. إذا يصبح لدينا من أجل التوزيع الفوق الهندسي:

$$E(X)=\frac{KM}{N}$$

بحساب مماثل، لكن أطول نسبيا، نتحصل على:

$$V(X)=\frac{KM}{N}\left(\frac{(N-M)(N-K)}{N(N-1)}\right)$$

نلاحظ هنا أن الطريقة لحساب $E(X)$ تمثلت في حساب المجموع من خلال كتابته على شكل توزيع فوق هندسي آخر بمعالم مختلفة.

مثال 1.2.:

التوزيع فوق الهندسي له تطبيقات عديدة أهمها استعماله في حالة الرغبة في قبول أو عدم قبول عينة ما، المثال التالي يوضح ذلك. بائع تجزئة يشتري سلع على شكل مجموعات حيث كل سلعة يمكن أن تقبل أو تكون فاسدة (غير مقبولة). ليكن:

N : عدد السلع في كل مجموعة

M: عدد السلع الفاسدة (غير صالحة) في كل مجموعة.

إذا يمكننا حساب احتمال أن تحوي مجموعة حجمها K على x سلعة فاسدة. لنكون أكثر دقة، نفرض أن التاجر استلم مجموعة من 25 جزء من آلة معينة أين يعتبر جزء منها مقبول فقط إذا استجاب لعتبة معينة. اخترنا عشر أجزاء (10) ووجدنا أنها كلها مقبولة. ما هو احتمال تحقق هذا الحادث إذا علمنا أن هناك 6 أجزاء فاسدة في المجموعة؟ بتطبيق التوزيع فوق الهندسي مع

$X \sim H.G(N=25, M=6, K=10)$ نتحصل على:

$$P(X=0)=\frac{C_6^0 C_{19}^{10}}{C_{25}^{10}}=0{,}028$$

يظهر من هذا أن هناك احتمال ضئيل لتحقق هذا الحادث.

2.ت. توزيع ثنائي الحد *Binomial Distribution*:

توزيع ثنائي الحد، أحد التوزيعات المتقطعة الأكثر استعمالا، حيث يعتمد على فكرة تحقيق وتكرار تجربة برنولي. هذه الأخيرة (نسبة إلى *James Bernoulli*، أحد مؤسسي نظرية الاحتمالات) تمثل كل تجربة عشوائية نتائجها تأخذ فقط حالتين. نقول عن متغير عشوائي X يتبع توزيع برنولي، $X \sim B(p)$ إذا كان:

$$X=\begin{cases}1 \text{ باحتمال } p\\ 0 \text{ باحتمال } 1-p\end{cases} \quad 0\le p\le 1 \;\ldots\ldots (2.3)$$

القيمة $X=1$ غالبا ترمز إلى حادث "نجاح" أو "تحقق" و p ترمز إلى احتمال تحققه. القيمة $X=0$ تمثل الحادث "فشل" أو "عدم تحقق" و $(1-\mathrm{p})$ ترمز إلى احتمال حدوثه. يعطى المتوسط (الأمل الرياضي) والتباين لتوزيع برنولي ب:

$$E(X)=1p+0(1-p)=p$$
$$V(X)=(1-p)^2p+(0-p)^2(1-p)=p(1-p)$$

العديد من التجارب العشوائية يمكن نمذجتها بتوزيع برنولي، أبسطها هي رمي قطعة نقدية مرة واحدة حيث نرمز ب" $\mathrm{X}=1$" إذا أظهرت القطعة الوجه و p يمثل احتمال الحصول على الوجه. أمثلة أخرى، انتخابات ($X=1$ إذا تحصل المترشح على الصوت) و حدوث مرض (p احتمال أن يصاب شخص عشوائي بالعدوى).

إذا تتابعت n تجربة عشوائية لبرنولي، نعرف الحوادث:

$$A_i=\{X=1، \text{ في التجربة } i\}, i=1,2,\ldots.,n$$

بافتراض أن الحوادث A_1،.....،A_n هي مجموعة مستقلة من الحوادث (كما هو الحال في تجربة رمي قطعة نقدية)، يكون من السهل استنتاج توزيع عدد مرات "النجاح" من n تجربة. نعرف المتغير Y ب:

الحادث $\{Y=y\}$ يتحقق فقط إذا تحقق من الحوادث A_1،.....A_n بالضبط y حادث وبالضرورة $n-y$ منها لا يتحقق. حالة خاصة من بين حالات التحقق ل n تجربة عشوائية مستقلة لبرنولي هي $A_1 \cap A_2 \cap A_3^c \cap \dots \cap A_{n-1} \cap A_n^c$. هذه الحالة لها احتمال تحقق:

$$P\left(A_1 \bigcap A_2 \bigcap A_3^c \bigcap \dots \bigcap A_{n-1} \bigcap A_n^c\right) = pp(1-p) \dots p(1-p)$$
$$= p^y(1-p)^{n-y}$$

أين استعملنا خاصية استقلال الحوادث في حساب الاحتمال. نشير هنا إلى أن حساب الاحتمال غير مرتبط بتحقق مجموعة محددة تحوي y حادث ولكن بأي مجموعة تحوي y حادث. بوضع كل هذا مع بعض، نلاحظ أن تتابع ل n حادث مع تحقق y حادث بالضبط له احتمال حدوث $p^y(1-p)^{n-y}$ وبما أنه لدينا C_n^y تتابع ممكن محققة ل y حادث فإنه يكون لدينا:

$$P(Y=y/n,p) = C_n^y p^y (1-p)^{n-y}, y = 0,1,2, \dots, n$$

يسمى Y متغير عشوائي ثنائي الحد ونكتب $Y \rightsquigarrow B(n,p)$.

المتغير العشوائي Y يمكن أن يعرف بطريقة مغايرة و مكافئة على النحو التالي: ليكن لدينا تتابع n تجربة عشوائية، مستقلة ومتماثلة لبرنولي كل منها لها احتمال p لتحقق الحادث "نجاح". نعرف المتغيرات العشوائية $X_1, \dots \dots, X_n$ ب:

$$X_i = \begin{cases} 1 & \text{باحتمال } p \\ 0 & \text{باحتمال } 1-p \end{cases} \qquad 0 \le p \le 1$$

إذا المتغير العشوائي:

$$Y = \sum_{i=1}^{n} X_i$$

يخضع لتوزيع ثنائي الحد $B(n,p)$,

الخاصية:

$$\sum_{y=0}^{n} P(Y=y) = \sum_{y=0}^{n} C_n^y p^y (1-p)^{n-y} = 1$$

تبرهن من النظرية العامة التالية.

نظرية 1.2. (**نظرية ثنائي الحد**):

من أجل كل عدد حقيقي x و y وكل عدد طبيعي $n \ge 0$، لدينا:

$$(x+y)^n = \sum_{i=0}^{n} C_n^i x^i y^{n-i} \dots \dots \dots (2.4)$$

البرهان:

نكتب: $(x+y)^n=(x+y)(x+y)\dots(x+y)$

ونبحث عن كيفية الحصول على الطرف الأيمن من المعادلة (4.2). من أجل كل معامل $(x+y)$ نختار سواء x أو y، ونضرب الخيارات n معا. من أجل $i=0,1,\dots,n$ ، عدد المرات التي يظهر بها، مثلا x، بالضبط i مرة هو C_n^i. وبالضرورة y يظهر $n-i$ مرة. إذا، هذا الطرف هو من الشكل $C_n^i x^i y^{n-i}$ وفي النهاية نتحصل على العلاقة المرجوة (4.2).

بأخذ $x=p$ و $y=1-p$ في العلاقة (4.2) نتحصل على:

$$(p+(1-p))^n=\sum_{i=0}^{n} C_n^i p^i (1-p)^{n-i}=1$$

حيث نلاحظ أن كل طرف في المجموع هو احتمال ثنائي الحد وهو برهان الخاصية أعلاه.

حالة خاصة أخرى يمكن استنتاجها من العلاقة (4.2) هي لما نأخذ $x=y=1$ ونتحصل على العلاقة:

$$\sum_{i=0}^{n} C_n^i=2^n$$

المتوسط أو الأمل الرياضي لمتغير ثنائي الحد يعطى ب:

$$E(X)=\sum_{x=0}^{n} xC_n^x p^x(1-p)^{n-x}=\sum_{x=1}^{n} xC_n^x p^x(1-p)^{n-x}$$

(الطرف الموافق ل $x=0$ هو معدوم). باستعمال العلاقة $xC_n^x=nC_{n-1}^{x-1}$ نتحصل على:

$$E(X)=\sum_{x=1}^{n} nC_{n-1}^{x-1} p^x(1-p)^{n-x}$$

$$E(X)=\sum_{y=0}^{n-1} nC_{n-1}^{y} p^{y+1}(1-p)^{n-(y+1)} \quad \text{بوضع } y=x-1$$

$$E(X)=np\sum_{y=0}^{n-1} nC_{n-1}^{y} p^{y}(1-p)^{n-1-y}$$

$$E(X)=np\,,$$

المجموع الأخير يساوي 1 كونه مجموع أطراف متغير عشوائي يخضع لتوزيع ثنائي الحد $Y\sim B(n-1,p)$.

لحساب التباين $V(X)$ نحسب أولا $E(X^2)$. لدينا:

$$E(X^2)=\sum_{x=0}^{n}x^2C_n^xp^x(1-p)^{n-x}$$

لحساب هذا المجموع، نبدأ أولا ببعض التحويلات، نكتب:

$$x^2C_n^x=x\frac{n!}{(x-1)!\,(n-x)!}=xnC_{n-1}^{x-1}$$

بالتعويض في $E(X^2)$ نتحصل على:

$$E(X^2)=n\sum_{x=0}^{n}xC_{n-1}^{x-1}p^x(1-p)^{n-x}$$

$$E(X^2)=n\sum_{y=0}^{n-1}(y+1)C_{n-1}^{y}p^{y+1}(1-p)^{n-(y+1)}$$ بوضع $y=x-1$

$$E(X^2)=np\sum_{y=0}^{n-1}yC_{n-1}^{y}p^{y+1}(1-p)^{n-1-y}+np\sum_{y=0}^{n-1}C_{n-1}^{y}p^{y+1}(1-p)^{n-1-y}$$

من الواضح الآن أن المجموع الأول من اليسار يساوي $(n-1)p$ كونه الأمل الرياضي لمتغير $Y \sim B(n-1,p)$ بينما المجموع الثاني يساوي 1. ومنه:

$$E(X^2)=n(n-1)p^2+np$$

ونتحصل في النهاية على:

$$V(X)=\ n(n-1)p^2+np-(np)^2$$
$$V(X)=np(1-p).$$

مثال 2.2.:

نفرض أننا نهتم بإيجاد أو حساب احتمال ظهور على الأقل مرة واحدة الرقم 6 من رمي 4 مرات متتالية زهرة نرد متزنة. هذه التجربة العشوائية يمكن نمذجتها بتتابع برنولي باحتمال نجاح $p=\frac{1}{6}$. نعرف المتغير العشوائي X ب:

X : عدد مرات الحصول على الرقم 6 من التجربة العشوائية و نكتب $B(4,1/6)$ $X \sim$ ومنه

$$P(X>0)=1-P(X=0)=1-C_4^0\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^4=0{,}518$$

الآن نقوم بتجربة أخرى، نرمي زوج من زهرات النرد 24 مرة ونتساءل عن احتمال الحصول على الأقل على زوج من الرقم 6. مرة أخرى يمكن نمذجة هذه التجربة العشوائية بتوزيع ثنائي الحد أو تتابع برنولي بحيث $p=P(\text{زوج من } 6)=\frac{1}{36}$.

نعرف المتغير العشوائي Y ب:

Y : عدد مرات الحصول على زوج من الرقم 6 من التجربة العشوائية و نكتب $Y \leadsto B(24, 1/36)$ ومنه

$$P(Y > 0) = 1 - P(Y = 0) = 1 - C_{24}^{0}\left(\frac{1}{36}\right)^{0}\left(\frac{35}{36}\right)^{24} = 0{,}491.$$

2.ث. توزيع بواسون (*Poisson Distribution*):

توزيع بواسون هو توزيع واسع الاستعمال من بين التوزيعات المتقطعة ويمكن استعماله كنموذج للعديد من التجارب المختلفة. مثلا عند نمذجة الظواهر التي من خلالها ننتظر حدوث أو ظهور حادث (قدوم حافلة، قدوم زبون إلى البنك...)، عدد مرات ظهور الحادث خلال مجال زمني يمكن أحيانا نمذجته بتوزيع "بواسون". واحدة من بين أهم الفرضيات المبني عليها توزيع بواسون هي أنه من أجل مجال أو فترة زمنية صغيرة احتمال حدوث أو ظهور الحادث هو متناسب مع طول فترة الانتظار الزمنية الأمر الذي يجعل من المنطقي استعمال توزيع بواسون لنمذجة الحالات والامثلة المذكورة، حيث من البديهي، مثلا، فرض أنه كلما زادت فترة انتظارنا كلما زاد احتمال مجيء زبون إلى البنك.

لدى توزيع بواسون معلمة واحدة λ، تسمى أحيانا **بمعلمة الكثافة**. نقول عن متغير عشوائي X، الذي يأخذ القيم الطبيعية، أنه يخضع لتوزيع بواسون، ونكتب $X \leadsto \mathrm{P}(\lambda)$، إذا كان:

$$P\big(X = x/\lambda\big) = \frac{e^{-\lambda}\lambda^{x}}{x!}, x = 0,1, \dots \qquad (2.5)$$

لتبيان أن $\sum_{x=0}^{\infty} P\big(X = x/\lambda\big) = 1$، نذكر أولا بصيغة تايلور ''Taylor'' ل e^{x}،

$$e^{x} = \sum_{i=0}^{\infty} \frac{x^{i}}{i!}$$

ومنه

$$\sum_{x=0}^{\infty} P\big(X = x/\lambda\big) = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^{x}}{x!} = e^{-\lambda}e^{\lambda} = 1$$

الأمل الرياضي ل X يمكن حسابه بسهولة:

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^{x}}{x!} \\ &= \sum_{x=1}^{\infty} x\frac{e^{-\lambda}\lambda^{x}}{x!} \end{aligned}$$

$$= \lambda e^{-\lambda} \sum_{1=0}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}$$

$$= \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^{y}}{y!} \qquad \text{بوضع} \qquad (y = x - 1)$$

$$= \lambda$$

بحساب مماثل نتحصل على:

$$V(X) = \lambda$$

إذا المعلمة λ هي متوسط وتباين توزيع بواسون.

مثال 3.2. (**فترة الانتظار**):

كمثال عن فترة الانتظار نأخذ مستقبل مكالمات هاتفية الذي يستقبل في المتوسط 5 مكالمات كل 3 دقائق. ما هو احتمال ان لا يستقبل أي مكالمة في الدقيقة القادمة؟ على الأقل مكالمتين؟

نرمز ب X لعدد المكالمات في الدقيقة، إذا X يخضع لتوزيع بواسون، $X \sim P(\lambda = 5/3)$، مع $E(X) = \lambda = \frac{5}{3}$. إذا:

$$P(X = 0) = \frac{e^{-\frac{5}{3}}\left(\frac{5}{3}\right)^{0}}{0!} = e^{-\frac{5}{3}} = 0{,}189$$

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1)$$

$$= 1 - 0{,}189 - \frac{e^{-\frac{5}{3}}\left(\frac{5}{3}\right)^{1}}{1!}$$

$$= 0{,}496$$

نشير إلى أنه بالإمكان حساب احتمالات توزيع بواسون بسرعة باستعمال العلاقة التراجعية التالية:

$$P(X = x) = \frac{\lambda}{x} P(X = x - 1), \qquad x = 1,2, \dots \dots (2.6)$$

يبرهن ببساطة على هذه العلاقة بالرجوع إلى دالة الكثافة الاحتمالية لبواسون (5.2).

علاقات تراجعية أخرى يمكن وضعها للتوزيعات المتقطعة. مثلا، إذا كان $Y \sim B(n, p)$ فإنه يمكن كتابة:

$$P(Y = y) = \frac{(n - y + 1)}{y} \frac{p}{1 - p} P(Y = y - 1) \dots \dots (2.7)$$

العلاقتين (6.2) و (7.2) يمكن استعمالهما لوضع الصيغة التقريبية لتوزيع ثنائي الحد بتوزيع بواسون.

أولا، بوضع $\lambda = np$ ولما تكون p صغيرة فإنه يمكننا كتابة:

$$\frac{(n-y+1)}{y}\frac{p}{1-p} = \frac{np-p(y-1)}{y-py} = \frac{\lambda}{y}$$

حيث الطرفين $p(y-1)$ و py يمكن إهمالهما (p صغيرة). إذا عند هذا المستوى من التقريب، العلاقة (7.2) تصبح:

$$P(Y=y) = \frac{\lambda}{y}P(Y=y-1) \ldots\ldots (2.8)$$

التي تمثل العلاقة التراجعية لتوزيع بواسون (6.2).

ثانيا، لإكمال المقاربة، وبما أن باقي الاحتمالات تستنتج من العلاقة (8.2)، يبقى فقط إثبات أن $P(X=0) = P(Y=0)$. الآن:

$$P(Y=0) = (1-p)^n = \left(1-\frac{np}{n}\right)^n = \left(1-\frac{\lambda}{n}\right)^n$$

رمزنا من قبل ب $np = \lambda$. لدينا أيضا من أجل λ ثابتة العلاقة

$$\lim_{n\to+\infty}\left(1-(\lambda/n)\right)^n = e^{-\lambda}$$

إذا من أجل قيمة كبيرة ل n يكون لدينا التقريب:

$$P(Y=0) = \left(1-\frac{\lambda}{n}\right)^n \approx e^{-\lambda} = P(X=0),$$

الأمر الذي يتم تقريب توزيع ثنائي الحد بتوزيع بواسون. نذكر بأن التقريب يكون أكثر دقة لما تكون n كبيرة و p صغيرة، الأمر الأكثر اعتياديا وحدوثا. هذا التقريب يسمح لنا أيضا بحساب معاملات ثنائيات الحد والقوى من أجل n كبيرة.

مثال 4.2. (**تقريب بواسون لتوزيع ثنائي الحد**):

كاتب عمومي يخطئ في المتوسط أثناء كتابته مرة واحدة كل 500 كلمة. الصفحة المعيارية تحوي 300 كلمة. ما هو احتمال أن لا يكون هناك أكثر من خطأين في 5 صفحات؟

بافتراض أن كتابة كلمة هي تسلسل برنولي مع احتمال "النجاح" $P = 1/500$ (لاحظ أننا أشرنا بالنجاح للخطأ في الكتابة) و أن الكلمات كتاباتها مستقلة، إذا X: عدد الأخطاء في 5 صفحات (1500 كلمة) يتبع توزيع ثنائي الحد $B(1500, 1/500)$، ومنه:

$$P(X \le 2) = \sum_{x=0}^{2} C_{1500}^{x}\left(\frac{1}{1500}\right)^x\left(\frac{499}{1500}\right)^{1500-x}$$

حيث نلاحظ صعوبة الحساب. لو نستعمل تقريب بواسون مع $\left(\lambda = np = 1500(1/500) = 3\right)$ نتحصل على:

$$P(X \le 2) = e^{-3}\left(1 + 3 + \frac{3^2}{2}\right) = 0,4232.$$

2.ج. توزيع ثنائي الحد السالب ***Negative Binomial Distribution***:

توزيع ثنائي الحد المشار إليه سابقا يهتم بحساب عدد مرات النجاح خلال عدد محدد من تجارب برنولي. لكن لو نهتم في المقابل بعدد مرات تكرار تجربة برنولي الضرورية للحصول على عدد محدد من مرات النجاح، نتحصل على توزيع ثنائي الحد السالب.

في تتابع تجارب مستقلة لبرنولي، ليكن المتغير X الذي يشير إلى عدد التجارب اللازمة أو الضرورية التي من أجلها يتحقق r حالة نجاح ، بحيث r هو رقم صحيح موجب. يكون إذا:

$$P(X = x/r, p) = C_{x-1}^{r-1} p^r (1-p)^{x-r}, x = r, r+1, \dots \dots (2.9)$$

ونقول أن المتغير X يتبع أو يخضع لتوزيع ثنائي الحد السالب ونكتب $X \sim NB(r,p)$.

العلاقة (9.2) تنتج من توزيع ثنائي الحد. الحادث $\{X = x\}$ يمكن أن يتحقق إذا وفقط إذا تحقق بالضبط $r-1$ نجاح في $x-1$ محاولة الأولى و نجاح في المحاولة رقم x. احتمال تحقق $r-1$ نجاح في $x-1$ محاولة أو تجربة هو احتمال ثنائي الحد $C_{x-1}^{r-1} p^{r-1} (1-p)^{x-r}$، ومع احتمال p أن يتحقق نجاح في التجربة رقم x. بضرب هذه الاحتمالات نتحصل على صيغة دالة الكثافة الاحتمالية لتوزيع ثنائي الحد السالب (9.2).

توزيع ثنائي الحد السالب يعرف أحيانا بدلالة المتغير العشوائي Y: عدد مرات الفشل (عدم تحقق) قبل تحقق r حالة نجاح. إحصائيا هذا التعريف مماثل للتعريف السابق المعطى أعلاه بدلالة المتغير X. من خلال تعريف المتغيرين X و Y يكون لدينا: $Y = X - r$.

باستعمال العلاقة بين X و Y، الصيغة البديلة لتوزيع ثنائي الحد السالب تعطى ب:

$$P(Y = y/r, p) = C_y^{r+y-1} p^r (1-p)^y, y = 0,1, \dots \dots (2.10).$$

ماعدا ذكر ذلك، عند إشارتنا لاحقا إلى توزيع ثنائي الحد السالب $NB(r,p)$ فإننا نقصد الصيغة (10.2).

في الواقع توزيع ثنائي الحد السالب أخذ تسميته من العلاقة:

$$C_y^{r+y-1} = (-1)^y C_y^{-r} = (-1)^y \frac{(-r)(-r-1)(-r-2)\dots\dots(-r-y+1)}{(y)(y-1)(y-2)\dots\dots(2)(1)}$$

التي هي في الواقع تعريف لمعادلة معاملات ثنائي الحد ذو قيم صحيحة سالبة. بتعويضها في العلاقة (10.2) نتحصل على:

$$P(Y = y) = (-1)^y C_y^{-r} p^r (1-p)^y, y = 0,1, \dots$$

الصيغة التي تشابه صيغة توزيع ثنائي الحد.

نشير أنه من الصعب التحقق بالنسبة لتوزيع ثنائي الحد السالب من أن $\sum_{y=0}^{\infty} P(Y=y)=1$، لكنها تتبع من توسعة لنظرية ثنائي الحد، توسعة تشمل المعاملات السالبة. عرض مفصل يمكن الاطلاع عليه لدى $Feller$ (1968).

متوسط وتباين Y يمكن حسابهما باستعمال الطرق نفسها عند حسابهما من أجل توزيع ثنائي الحد:

$$E(Y)=\sum_{y=0}^{\infty} yC_y^{r+y-1}p^r(1-p)^y$$

$$=\sum_{y=0}^{\infty}\frac{(r+y-1)!}{(y-1)!\,(r-1)!}p^r(1-p)^y$$

$$=\sum_{y=1}^{\infty} rC_{y-1}^{r+y-1}p^r(1-p)^y$$

الآن نضع $z=y-1$ والمجموع يصبح:

$$E(Y)=\sum_{z=0}^{\infty} rC_z^{r+z}p^r(1-p)^{z+1}$$

$$=r\frac{1-p}{p}\sum_{z=0}^{\infty} C_z^{(r+1)+z-1}p^r(1-p)^z$$

حيث مجموع عناصر توزيع ثنائي الحد السالب $(r+1,p)$ تساوي الواحد، نتحصل في النهاية على:

$$E(Y)=\frac{r(1-p)}{p}$$

بحساب مماثل نتحصل على:

$$V(Y)=\frac{r(1-p)}{p^2}$$

توجد أهمية وإفادة في إعادة صياغة معلمة توزيع ثنائي الحد السالب بدلالة متوسطه. بتعريف المعلمة $\mu=r\,(1-p)/p$ ، نجد أن $E(Y)=\mu$ وببعض العمليات الجبرية نجد أيضا:

$$V(Y)=\mu+\frac{1}{r}\mu^2$$

حيث يكون التباين دالة تربيعية بدلالة المتوسط. هذه العلاقة هي مهمة ومفيدة في تحليل البيانات ونظرية الاعتبارات (Morris 1982).

نشير إلى أن عائلة توزيع ثنائي الحد السالب تشمل توزيع بواسون كحالة نهاية. إذا كان $r\to\infty$ و $p\to 1$ بحيث $r(1-p)\to\lambda$، و $0<\lambda<\infty$ فإن:

$$E(Y) = \frac{r(1-p)}{p} \to \lambda$$

$$V(Y) = \frac{r(1-p)}{p^2} \to \lambda$$

اللذان يتوافقان مع متوسط وتباين توزيع بواسون. لبرهنة أن $NB(r,p) \to P(\lambda)$ يمكننا إظهار ذلك بأن جميع الاحتمالات تتقارب.

مثال 5.2. (معاينة ثنائي الحد العكسية):

تقنية معروفة ب " معاينة ثنائي الحد العكسية" "*Inverse binomial sampling*" هي مستعملة كثيرا في معاينة المجتمعات البيولوجية. إذا كانت نسبة الأفراد الذين يملكون خاصية معينة في مجتمع ما هي "p" وقمنا باختيار عينة، واحدة تلو الأخرى، حتى نتحصل على r فرد يملك هذه الخاصية، إذا عدد افراد (عناصر) العينة هو متغير عشوائي يخضع لقانون ثنائي الحد السالب.

مثلا، نفرض أننا نهتم في مجتمع الحشرات الطائرة بنسبة تلك اللواتي يملكن نوع معين من الأجنحة وقررنا لأجل ذلك اختيار عينة وتفحصها حتى نتحصل على 100 حشرة من هذا النوع. احتمال ان نتفحص على الأقل N حشرة طائرة هو (باستعمال العلاقة (9.2)):

$$P(X \geq N/r = 100, p) = \sum_{x=N}^{\infty} C_{x-1}^{99} p^{100} (1-p)^{x-100}$$

$$= 1 - \sum_{x=100}^{N-1} C_{x-1}^{99} p^{100} (1-p)^{x-100}$$

من أجل قيم ل p و N، يمكننا حساب هذه العبارة مع الإشارة على أنه من الأسهل استعمال العلاقة التراجعية لحساب ذلك.

هذا المثال يظهر أن توزيع ثنائي الحد السالب، مثل توزيع بواسون، يمكن استعماله لنمذجة الظواهر التي من خلالها ننتظر حدوث حادث معين. في هذه الحالة الحادث هو تحقق عدد محدد من حالات "النجاح".

2.ح.التوزيع الهندسي *Geometric Distribution*:

التوزيع الهندسي هو أبسط توزيع لظاهرة أو تجربة انتظار وهو يمثل حالة خاصة من توزيع ثنائي الحد السالب. بوضع $r = 1$ في العلاقة (9.2) نتحصل على:

$$P(X = x/p) = p(1-p)^{x-1}, x = 1,2, \dots \dots$$

هذه العلاقة تعرف دالة الكثافة الاحتمالية لمتغير X يخضع للتوزيع الهندسي مع احتمال "النجاح" p ونكتب $X \rightsquigarrow G(p)$. X يمكن أن يفسر على أنه رقم التجربة أو المحاولة حتى تظهر أول حالة "نجاح"، إذا نحن ننتظر ظهور أو تحقق "نجاح".

الخاصية، $\sum_{x=1}^{\infty} P(X = x) = 1$، تبرهن باستعمال خاصية المتتالية الهندسية. من أجل كل عدد حقيقي a مع $|a| < 1$، لدينا:

$$\sum_{x=1}^{\infty} a^{x-1} = \frac{1}{1-a}$$

متوسط وتباين متغير X يخضع للتوزيع الهندسي يمكن حسابه باستعمال الصيغ الموافقة لتوزيع ثنائي الحد السالب أعلاه و بكتابة $X = Y + 1$ نتحصل على:

$$E(X) = E(Y) + 1 = \frac{1}{p}$$

و

$$V(X) = \frac{1-p}{p^2}$$

التوزيع الهندسي له خاصية مميزة تعرف بخاصية "عدم الذاكرة". من أجل قيم صحيحة $s > t$، يتحقق:

$$P(X > s/X > t) = P(X > s - t), \dots \dots (2.11)$$

هذا يعني أن التوزيع الهندسي لا يسجل ولا يتذكر (إن صح التعبير) ماذا حدث. احتمال الحصول على "$s - t$ " حالة فشل إضافية مع العلم أننا شاهدنا t حالة فشل هو نفسه احتمال مشاهدة "$s - t$ " حالة فشل في بداية التجربة. بعبارة أخرى، احتمال الحصول على مجموعة من حالات الفشل هو فقط مرتبط بطول المجموعة وليس بتموضعها في الزمن.

لتوضيح العلاقة (11.2)، نشير أولا أنه من أجل كل عدد صحيح n، حسب التوزيع الهندسي، $P(X > n)$ يمثل احتمال ان لا يكون أي "نجاح" في n تجربة ومنه:

$$P(X > n) = (1 - p)^n \dots \dots (2.12)$$

البرهان:

لدينا

$$P(X > n) = \sum_{x=n+1}^{\infty} p(1-p)^{x-1} = p \sum_{x=n+1}^{\infty} (1-p)^{x-1}$$

$$= p \sum_{x=n+1}^{\infty} (1-p)^{x-1}$$

$$= p(1-p)^{n+1-1} \lim_{\alpha \to \infty} \frac{1 - (1-p)^{\alpha}}{1 - (1-p)}$$ مجموع متتالية هندسية

$$= (1-p)^n$$

ومنه

$$P(X > s/X > t) = \frac{P\left(X > s \text{ و } X > t\right)}{P(X > t)}$$

$$= \frac{P(X > s)}{P(X > t)} = \frac{(1-p)^s}{(1-p)^t}$$

$$= (1-p)^{s-t}$$
$$= P(X > s-t)$$

مثال 6.2. (**أوقات التعطل**):

يستعمل التوزيع الهندسي أحيانا لنمذجة "مدة الحياة" أو "المدة الزمنية حتى يحدث العطب" لآلة أو تركيب معين. مثلا، إذا كان احتمال ان يتعطل مصباح كهربائي في أي يوم هو $0{,}001$، إذا احتمال أن يعيش مصباح على الأقل 30 يوم هو:

$$P(X > 30) = \sum_{x=31}^{\infty} (0{,}001)(1-0{,}001)^{x-1} = (1-0{,}001)^{30} = (0{,}999)^{30} = 0{,}970.$$

غياب أو عدم الذاكرة للتوزيع الهندسي تمثل خاصية "عدم التقدم في العمر". تشير هذه الخاصية إلى أن التوزيع الهندسي لا يمك استعماله لنمذجة "مدة الحياة" عندما يكون احتمال "التعطل أو العطب" يزداد بزيادة العمر او الزمن. توجد توزيعات أخرى تستعمل في هذه الحالة.

3. التوزيعات الاحتمالية المستمرة ***Continuous Distribution***:

في هذا الجزء سوف نتطرق إلى أهم التوزيعات المستمرة الأمر الذي يعني أن هناك توزيعات إحصائية أخرى كثيرة لن نتطرق إليها. أكثر من ذلك يمكن إثبات ان كل دالة موجبة قابلة للتكامل يمكن تحويلها إلى دالة كثافة احتمالية (توزيع مستمر).

3.أ. توزيع أحادي الشكل المستمر ***Continuous Uniform Distribution***:

توزيع أحادي الشكل لمتغير عشوائي X يعرف بتوزيع كتلته بالتساوي على طول المجال $[a,b]$ ونكتب $X \backsim U([a,b])$. تعطى دالة كثافته الاحتمالية ب:

$$f(x/a,b)=\begin{cases}\dfrac{1}{b-a} & \text{إذا كان } x\in[a,b]\\ 0 & \text{الحالة العكسية}\end{cases} \quad \dots\dots(3.1)$$

من السهل التحقق أن $\int_a^b f(x)dx=1$.

من أجل أمله الرياضي أو تباينه لدينا:

$$E(X)=\int_a^b \frac{x}{b-a}dx$$
$$=\frac{b+a}{2}$$

$$V(X)=\int_a^b \frac{\left(x-\frac{b+a}{2}\right)^2}{b-a}dx$$
$$=\frac{1}{b-a}\int_a^b\left(x^2+\frac{(b+a)^2}{4}-(b+a)x\right)dx$$
$$=\frac{1}{b-a}\left[\frac{x^3}{3}+\frac{(b+a)^2}{4}x-\frac{(b+a)}{2}x^2\right]_a^b$$

ببعض الحسابات والتبسيطات نجد في النهاية:

$$V(X)=\frac{(b-a)^2}{2}$$

3.ب. توزيع غاما ***Gamma Distribution***:

عائلة توزيع غاما هي عائلة من التوزيعات المرنة المعرفة على $]0,\infty[$. إذا كان α عدد ثابت موجب، التكامل

$$\int_0^\infty t^{\alpha-1}e^{-t}dt$$

هو محدود (منتهي). إذا كانت α عدد صحيح موجب، التكامل يمكن أن يعبر عنه بصيغة مغلقة، الامر غير ممكن في الحالة المغايرة. في كل الحالات تعرف قيمه **دالة غاما** المعبر عنها ب:

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt \dots\dots (3.2)$$

تحقق دالة غاما العديد من الخواص أهمها:

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha)\,, \alpha > 0 \dots\dots (3.3)$$

التي يمكن التحقق منها بالتكامل بالتجزئة. بجمع الخاصية (3.3) مع البديهية $\Gamma(\alpha) = 1$، يكون لدينا من أجل كل عدد صحيح موجب n:

$$\Gamma(n) = (n-1)! \dots\dots (3.4)$$

(الخاصية المهمة الكثيرة الاستعمال ، التي سنراها لاحقا، هي $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$).

العبارتين (3.3) و (4.3) تعطيان صيغ تراجعية تسهل صعوبات حساب قيم دالة غاما التي تسمح لنا بحساب أي قيمة لدالة غاما من خلال معرفة فقط قيم $\Gamma(c)$، $0 < c \le 1$.

بما ان التكامل (2.3) هو موجب فإنه يتبع مباشرة أن الدالة:

$$f(t) = \frac{t^{\alpha-1}e^{-t}}{\Gamma(\alpha)}, 0 < t < \infty \dots\dots (3.5)$$

هي دالة كثافة احتمالية بحيث t هو متغير عشوائي يتبع توزيع غاما، ونكتب $T \leadsto G(\alpha)$.

في الواقع، عائلة دوال الكثافة الاحتمالية لغاما لها معلمتين (α, β)، ويمكن أن نتحصل على ذلك بالبحث عن دالة الكثافة الاحتمالية للمتغير العشوائي $X = \beta T$، حيث β هوعدد حقيقي ثابت موجب، من خلال تغيير متغير في العلاقة (5.3). نبين ذلك.

لدينا:

$$x = \beta t \Rightarrow t = \frac{x}{\beta} \text{ و } dt = \frac{dx}{\beta}$$

و لدينا من العلاقة (5.3) وبالتعويض يمكننا كتابة:

$$\int_0^\infty f(t)\,dt = 1 \Rightarrow \frac{1}{\Gamma(\alpha)}\int_0^\infty \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)} \frac{dx}{\beta} = 1$$

$$\Rightarrow \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^\infty x^{\alpha-1}e^{-(x/\beta)}\,dx = 1$$

(نشير إلى أن الجزء $\left(\int_0^\infty x^{\alpha-1}e^{-(x/\beta)}\,dx\right)$ يسمى **بالجزء الأساسي** من تكامل دالة الكثافة الاحتمالية ل $G(\alpha, \beta)$)

حيث يمكننا أن نستنتج صيغة دالة الكثافة الاحتمالية لعائلة غاما $G(\alpha,\beta)$ على النحو التالي:

$$f(x/\alpha,\beta)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{\alpha-1}e^{-(x/\beta)},\qquad 0<x<\infty,\qquad \alpha>0,\qquad \beta>0\dots\dots(3.6)$$

ونكتب $X \sim G(\alpha,\beta)$. تعرف المعلمة α بمعلمة **الشكل**، حيث تحدد شكل تطاول التوزيع، بينما تعرف المعلمة β بمعلمة **القياس** أو **السلم** كونها أكثر تأثيرها يكون على درجة تشتت وانتشار التوزيع الموافق.

متوسط أو أمل توزيع غاما $G(\alpha,\beta)$ يعطى ب:

$$E(X)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\int_0^{\infty}xx^{\alpha-1}e^{-x/\beta}\,dx$$

$$E(X)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\int_0^{\infty}x^{\alpha}e^{-x/\beta}\,dx.\dots\dots(3.7)$$

لحساب جزء التكامل من العلاقة (7.3)، نلاحظ أنه يمثل **الجزء الأساسي** من تكامل دالة الكثافة الاحتمالية لتوزيع $G(\alpha+1,\beta)$. أي من العلاقة (6.3) ومن أجل $\alpha,\beta>0$ يكون لدينا:

$$\int_0^{\infty}f(x/\alpha+1,\beta)=1\Rightarrow\frac{1}{\Gamma(\alpha+1)\beta^{\alpha+1}}\int_0^{\infty}x^{\alpha}e^{-(x/\beta)}dx=1$$

$$\Rightarrow\int_0^{\infty}x^{\alpha}e^{-(x/\beta)}dx=\Gamma(\alpha+1)\beta^{\alpha+1}$$

ومنه العلاقة (7.3) يصبح:

$$E(X)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\Gamma(\alpha+1)\beta^{\alpha+1}$$

$$=\alpha\beta$$

يحسب تباين توزيع غاما $G(\alpha,\beta)$ بطريقة مماثلة لطريقة حساب المتوسط. بالخصوص، عند حساب $E(X^2)$، نستعمل **الجزء الأساسي** من تكامل دالة الكثافة الاحتمالية لتوزيع $G(\alpha+2,\beta)$ حيث نجد:

$$E(X^2)=\alpha(\alpha+1)\beta^2\dots\dots(3.8)$$

ومنه:

$$V(X)=\alpha\beta^2.$$

مثال 1.3. (**العلاقة بين توزيعي غاما وبواسون**):

توجد علاقة مهمة بين توزيعي غاما وبواسون. إذا كان المتغير X يخضع لتوزيع غاما $G(\alpha,\beta)$، حيث α هو عدد طبيعي، فإنه من أجل كل x، يتحقق لدينا:

$$P(X \leq x) = P(Y \geq \alpha) \ldots\ldots (3.9)$$

بحيث $Y \sim P(x/\beta)$. العلاقة (9.3) يمكن توضيحها بسلسلة من التكاملات بالتجزئة على النحو التالي. بما أن α هو عدد طبيعي، نكتب $\Gamma(\alpha) = (\alpha-1)!$ ونتحصل على:

$$\begin{aligned} P(X \leq x) &= \frac{1}{(\alpha-1)!\,\beta^{\alpha}} \int_0^x t^{\alpha-1} e^{-t/\beta} dt \\ &= \frac{1}{(\alpha-1)!\,\beta^{\alpha}} \left[-t^{\alpha-1} \beta e^{-t/\beta} \Big|_0^x + \int_0^x (\alpha-1) t^{\alpha-2} \beta e^{-t/\beta} dt \right], \end{aligned}$$

أين استعملنا التكامل بالتجزئة بوضع $u = t^{\alpha-1}$ و $dv = e^{-t/\beta} dt$. نتابع:

$$\begin{aligned} P(X \leq x) &= \frac{-1}{(\alpha-1)!\,\beta^{\alpha-1}} x^{\alpha-1} e^{-x/\beta} + \frac{1}{(\alpha-2)!\,\beta^{\alpha-1}} \int_0^x t^{\alpha-2} e^{-t/\beta} dt \\ &= \frac{1}{(\alpha-2)!\,\beta^{\alpha-1}} \int_0^x t^{\alpha-2} e^{-t/\beta} dt - P(Y = \alpha-1) \text{ أين } Y \sim P(x/\beta) \\ &= \frac{1}{(\alpha-3)!\,\beta^{\alpha-2}} \int_0^x t^{\alpha-3} e^{-t/\beta} dt - [P(Y=\alpha-1) + P(Y=\alpha-2)] \end{aligned}$$

بالمواصلة بنفس الطريقة نتحصل على:

$$\begin{aligned} &= \frac{1}{(0)!\,\beta^{1}} \int_0^x t^{0} e^{-t/\beta} dt \\ &\quad - [P(Y=\alpha-1) + P(Y=\alpha-2) + \cdots + P(Y=1)] \\ &= 1 - [P(Y=\alpha-1) + P(Y=\alpha-2) + \cdots + P(Y=1) + P(Y=0)] \\ &= 1 - P(Y < \alpha) \\ &= P(Y \geq \alpha) \end{aligned}$$

أين تحصلنا على العلاقة (9.3).

توجد العديد من الحالات الخاصة لتوزيع غاما $G(\alpha,\beta)$. مثلا بوضع $\alpha = p/2$ و $\beta = 2$، بحيث p هو عدد طبيعي، تصبح دالة الكثافة الاحتمالية لتوزيع غاما (6.3) كالتالي:

$$f(x/\alpha,\beta) = \frac{1}{\Gamma(p/2) 2^{(p/2)}} x^{(p/2)-1} e^{-x/2}, \qquad 0 < x < \infty, \qquad p > 0, \quad \ldots\ldots (3.10)$$

التي تمثل دالة الكثافة الاحتمالية لتوزيع **كاي–تربيع** (***Chi-Squared Distribution***) بدرجة حرية p حيث نكتب $X \sim \chi_p^2$.

المتوسط والتباين لتوزيع كاي–تربيع يمكن حسابهما باستعمال علاقتي توزيع غاما فيكون:

$$E(X) = \alpha\beta = p$$

$$V(X) = \alpha\beta^2 = 2p$$

نشير إلى أن توزيع **كاي–تربيع** يلعب دور أساسي في الاستدلال الاحصائي عند المعاينة من مجتمع طبيعي. حالة خاصة أخرى مهمة لتوزيع غاما kتحصل عليها بوضع "$\alpha = 1$" فنجد:

$$f(x/\beta) = \frac{1}{\beta} e^{-x/\beta}, \qquad 0 < x < \infty, \qquad \dots \dots (3.11)$$

التي تمثل دالة الكثافة الاحتمالية للتوزيع الأسي ***(Exponential Distribution)*** ذو معلمة قياس β، ونكتب $X \sim E(\beta)$.

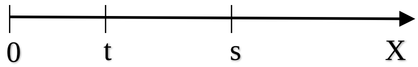
هنا كذلك متوسط وتباين التوزيع الأسي يمكن حسابهما باستعمال علاقتي توزيع غاما الموافقتين فيكون:

$$E(X) = \alpha\beta = \beta$$

$$V(X) = \alpha\beta^2 = \beta^2$$

التوزيع الأسي يمكن استعماله لنمذجة "**مدة الحياة**"، مثله مثلما يستعمل التوزيع الهندسي في الحالة المتقطعة. في الواقع، التوزيع الأسي يظهر ويتميز أيضا بخاصية "**عدم الذاكرة**"، تلك الخاصة بالتوزيع الهندسي.

إذا كان $X \sim E(\beta)$ حيث دالة كثافته الاحتمالية معطاة بالعلاقة (11.3)، إذا من أجل $s > t \geq 0$:



يكون لدينا:

$$P(X > s/X > t) = P(X > s - t)$$

لتوضيح ذلك، لدينا من جهة:

$$P(X > s/X > t) = \frac{P(X > s, X > t)}{P(X > t)}$$

$$= \frac{P(X > s)}{P(X > t)} \quad \text{بما أن } s > t$$

$$= \frac{\int_x^\infty \frac{1}{\beta} e^{-x/\beta}\, dx}{\int_x^\infty \frac{1}{\beta} e^{-x/\beta}\, dx}$$

$$= \frac{e^{-s/\beta}}{e^{-t/\beta}}$$

$$= e^{-(s-t)/\beta}$$

ومن جهة:

$$P(X > s - t) = \int_{(s-t)}^{\infty} \frac{1}{\beta} e^{-x/\beta}\, dx$$

$$= \left[-e^{-x/\beta}\right]_{(s-t)}^{\infty}$$

$$= e^{-(s-t)/\beta}$$

وهو المطلوب.

توزيع آخر مرتبط بعائلة غاما وعائلة التوزيعات الأسية هو توزيع "ويبل" (***Weibull Distribution***). إذا كان $X \sim E(\beta)$، إذا $Y = X^{1/\gamma}$ يتبع توزيع ويبل المعرف بالمعلمتين γ و β ، ونكتب $Y \sim W(\gamma, \beta)$، المعطاة دالة كثافته الاحتمالية ب:

$$f(y/\gamma, \beta) = \frac{\gamma}{\beta} y^{\gamma-1} e^{-y^{\gamma}/\beta}, \quad 0 < y < \infty, \quad \gamma > 0, \quad \beta > 0 \ldots\ldots.. (3.12)$$

نوضح هذه العلاقة. لدينا:

$$X \sim E(\beta) \Rightarrow P(X \leq x) = \int_0^x \frac{1}{\beta} e^{-x/\beta}\, dx \ldots\ldots (*)$$

و

$$y = x^{1/\gamma} \Rightarrow \begin{cases} x = y^{\gamma} \\ dx = \gamma y^{\gamma-1} dy \end{cases} \ldots\ldots (**)$$

بتعويض $(**)$ في $(*)$ نجد:

$$P(X \leq x) = \int_0^y \frac{1}{\beta} e^{-y^{\gamma}/\beta}\, \gamma y^{\gamma-1} dy$$

$$\Rightarrow f(y/\gamma, \beta) = \frac{\gamma}{\beta} y^{\gamma-1} e^{-y^{\gamma}/\beta}$$

وهو المطلوب (نشير أننا لم نغير المتغير داخل التكامل لكي يسهل تتبع التعويضات).

يتضح من العلاقة (12.3) أنه كان بالإمكان البداية بتوزيع ويبل ثم اشتقاق التوزيع الأسي كحالة خاصة منه بوضع ($\gamma = 1$). نشير في النهاية إلى أن توزيع ويبل يلعب دورا أساسيا في تحليل بيانات فترات التعطل كما هو أكثر إفادة في نمذجة الدوال العشوائية.

3.ت. التوزيع الطبيعي (*Normal Distribution*):

التوزيع الطبيعي (أحيانا يسمى بتوزيع غوس ***Gauss Distribution***) يلعب دورا أساسيا ومحوريا في نظرية الإحصاء. يوجد ثلاثة أسباب لهذا. أولا، التوزيع الطبيعي والتوزيعات المشتقة أو المرتبطة به هي سهلة التحليل (حتى وإن كان هذا لا يظهر للوهلة الأولى). ثانيا، التوزيع الطبيعي له شكل بياني متناظر، على صفة جرس، الأمر الذي يجعل منه يتطابق مع أغلب توزيعات المجتمعات الإحصائية حيث على الرغم من وجود توزيعات أخرى متناظرة إلا أن التعامل معها حسابيا هو أكثر تعقيدا منه بالنسبة للتوزيع الطبيعي. ثالثا، نظرية النهايات المركزية التي تشير إلى أنه تحت شروط وقيود معينة، التوزيع الطبيعي يمكن أن يستعمل كتقريب (توزيع تقريبي) لمجموعة متعددة من التوزيعات في العينات الإحصائية.

التوزيع الطبيعي له معلمتين، يرمز لهما غالبا ب μ و σ^2 اللتان تمثلان على التوالي المتوسط (الأمل الرياضي) والتباين. دالة الكثافة الاحتمالية للتوزيع الطبيعي لمتغير إحصائي X من أجل متوسط μ وتباين σ^2 تعطى ب:

$$f(x/\,\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/(2\sigma^2)},\qquad -\infty<x<\infty\ \dots\dots\dots(3.13)$$

ونكتب $X \rightsquigarrow N(\mu,\sigma^2)$.

إذا كان X متغير طبيعي، $X \rightsquigarrow N(\mu,\sigma^2)$، إذا المتغير العشوائي $Z=(X-\mu)/\sigma$ له أيضا توزيع طبيعي $N(0,1)$ يعرف **بالتوزيع الطبيعي المعياري** (أو فقط التوزيع المعياري) ونكتب $Z \rightsquigarrow N(0,1)$. يمكن اثبات ذلك بكتابة:

$$\begin{aligned}P(Z\le z)&=P\left(\frac{X-\mu}{\sigma}\le z\right)\\&=P(X\le z\sigma+\mu)\\&=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{z\sigma+\mu}e^{-(x-\mu)^2/(2\sigma^2)}\,dx\\&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-t^2/2}\,dt\quad \text{بتعويض}\ t=\frac{x-\mu}{\sigma}\end{aligned}$$

التي تظهر أن $P(Z\le z)$ هي دالة توزيع تراكمية لتوزيع طبيعي بمتوسط 0 وتباين 1.

يتبع من هذا أن كل احتمالات التوزيع الطبيعي يمكن حسابها من خلال التوزيع الطبيعي المعياري. أيضا، حسابات الأمل الرياضي يمكن تبسيطها باستعمال العلاقة التي تربط بين التوزيع الطبيعي والتوزيع الطبيعي المعياري.

مثلا، إذا كان $Z \rightsquigarrow N(0,1)$:

$$E(Z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}ze^{-z^2/2}\,dz=\ \frac{1}{\sqrt{2\pi}}e^{-z^2/2}\bigg|_{-\infty}^{\infty}=0$$

$$E(Z^2)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}z^2e^{-z^2/2}\,dz=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}zze^{-z^2/2}\,dz$$

$$= \frac{1}{\sqrt{2\pi}}\left[-ze^{-z^2/2}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-z^2/2}\,dz\right]$$
$$= \frac{1}{\sqrt{2\pi}}\left[0 + \sqrt{2\pi}\right] = 1$$

وكان $X \sim N(\mu, \sigma^2)$، يكون لدينا:

$$E(X) = E(Z\sigma + \mu) = \sigma E(Z) + \mu = \mu$$

كذلك وبطريقة مماثلة نجد:

$$V(X) = V(Z\sigma + \mu) = \sigma^2 V(Z) = \sigma^2$$

لإظهار أن تكامل دالة الكثافة الاحتمالية للتوزيع الطبيعي (13.3) لمجموعة التعريف يساوي 1 يكفي أن نبين ذلك من أجل دالة الكثافة الاحتمالية للتوزيع الطبيعي المعياري:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-z^2/2}\,dz = 1$$

نذكر أن هذا التوزيع المعياري متناظر بالنسبة ل 0 الأمر الذي يستلزم بأن يكون التكامل من $]-\infty; 0]$ مساوي للتكامل من $[0, \infty[$. في هذه الحالة تختزل الإشكالية إلى إثبات أن:

$$\int_0^{\infty} e^{-z^2/2}\,dz = \sqrt{\frac{\pi}{2}} \quad \ldots\ldots (3.14)$$

نذكر كذلك أن الدالة $\left(e^{-z^2/2}\right)$ ليس لها دالة أصلية كونها لا يمكن كتابتها على شكل دالة أساسية او مجموع دوال أساسية (هذا في الصيغة المغلقة)، أي أنه لا يمكننا حساب مباشرة التكامل (14.3). طريقة أخرى، بما أن طرفي العلاقة (14.3) هما موجبين، فإن إثباتها يكون مكافئ لإثبات تساوي مربعي طرفيها. بتربيع العلاقة (14.3) نتحصل على:

$$\left(\int_0^{\infty} e^{-z^2/2}\,dz\right)^2 = \left(\int_0^{\infty} e^{-t^2/2}\,dt\right)\left(\int_0^{\infty} e^{-u^2/2}\,du\right)$$
$$= \int_0^{\infty}\int_0^{\infty} e^{-(t^2+u^2)/2}\,dt\,du$$

بالانتقال إلى الإحداثيات القطبية. نضع:

$$t = r\cos\theta \quad \text{و} \quad u = r\sin\theta$$

يكون لدينا $t^2 + u^2 = r^2$ و $dtdu = rd\theta dr$ ومجال التكامل يصبح $0 < r < \infty$ و $0 < \theta < \pi/2$ (الحد الأعلى ل θ هو $\pi/2$ لأن u و t هما موجبين شرطا). يصبح لدينا الآن:

$$\int_0^{\infty}\int_0^{\infty} e^{-(t^2+u^2)/2}\,dt\,du = \int_0^{\infty}\int_0^{\pi/2} re^{-r^2/2}\,d\theta\,dr$$

$$= \frac{\pi}{2}\int_0^{\infty} re^{-r^2/2}\,dr$$
$$= \frac{\pi}{2}\left[-e^{-r^2/2}\right]_0^{\infty}$$
$$= \frac{\pi}{2}$$

ومنه

$$\int_0^{\infty} e^{-z^2/2}\,dz = \sqrt{\frac{\pi}{2}}$$

النتيجة التي تثبت العلاقة (14.3).

هذا التكامل هو مرتبط بدالة غاما. بوضع التغيير $w = \frac{1}{2}z^2$ في العلاقة (14.3) وبالتعويض الصحيح فيها يصبح :

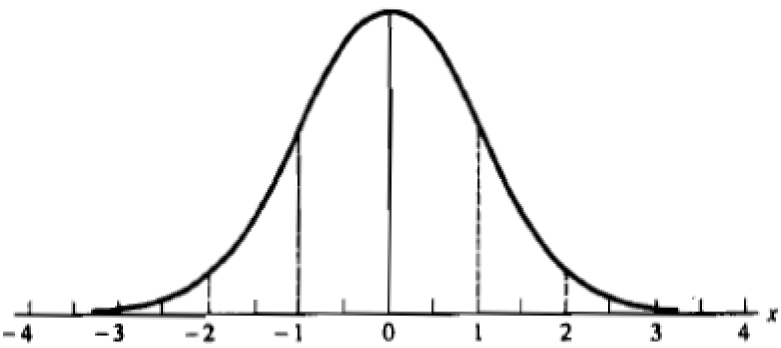
$$\int_0^{\infty} e^{-z^2/2}\,dz = \sqrt{\frac{\pi}{2}}$$
$$\Rightarrow \frac{1}{\sqrt{2}}\int_0^{\infty} w^{1-\frac{1}{2}}e^{-w/2}\,dw = \sqrt{\frac{\pi}{2}}$$
$$\Rightarrow \int_0^{\infty} w^{1-\frac{1}{2}}e^{-w/2}\,dw = \sqrt{\pi} = \Gamma\left(\frac{1}{2}\right) \dots\dots\dots (3.15)$$

التوزيع الطبيعي هو نوعا ما توزيع خاص أو حالة خاصة كون معلمتيه، المتوسط (μ) و التباين (σ^2)، يعطياننا معلومات كاملة عن شكل و موضع التوزيع. هذه الخاصية، أي ان التوزيع هو محدد ب μ و σ^2، هي ليست خاصة فقط بدالة الكثافة الاحتمالية للتوزيع الطبيعي بل حالة مشتركة لمجموعة أو عائلة من دوال الكثافة تسمى عائلات توزيع ***الموضع–القياس*** التي سنشير إليها لاحقا.

حسابات مباشرة سوف تظهر لنا أن دالة الكثافة الاحتمالية للتوزيع الطبيعي (13.3) تصل إلى قيمتها الأعظمية (المشتقة الأول تنعدم والمشتقة الثانية موجبة) عند $(x = \mu)$ ونقاط انعطافها (المشتقة الثانية تنعدم) عند $(x = \mu \pm \sigma)$ (عندهما المنحنى الموافق ينتقل من الشكل المقعر إلى الشكل المحدب). كذلك، احتمال احتواء 1، 2 أو 3 من الانحراف المعياري للمتوسط هو على التوالي:

$$P(|X-\mu| \le \sigma) = P(|Z| \le 1) = 0{,}6826$$
$$P(|X-\mu| \le 2\sigma) = P(|Z| \le 2) = 0{,}9544$$
$$P(|X-\mu| \le 3\sigma) = P(|Z| \le 3) = 0{,}9974$$

أين $X \backsim N(\mu, \sigma^2)$ و $Z \backsim N(0,1)$ والقيم الرقمية للاحتمالات يمكن الحصول عليها من العديد من برامج الإعلام الآلي أو من الجداول الإحصائية. غالبا قيم الأطراف الثنائية المستعملة هي 0,68، 0,95 و 0,99 على التوالي. على الرغم من أن هذه الأخيرة لا تمثل قيم مقربة لكنها جرت أكثر العادة على استعمالها. الشكل (1.3) أسفله يعرض دالة الكثافة الاحتمالية للتوزيع الطبيعي المعياري مع هذه القيم الأساسية.



الشكل 1.3. دالة كثافة التوزيع الطبيعي المعياري

من بين الاستعمالات العديدة للتوزيع الطبيعي، ومن بين أهمها، يستعمل لتقريب بعض التوزيعات الأخرى (هذا التقريب مبرر جزئيا بنظرية النهايات المركزية). مثلا، إذا كان X يخضع لتوزيع ثنائي الحد، $X \sim B(n,p)$، فإن $E(X) = np$ و $V(X) = np(1-p)$، حيث تحت شروط معينة توزيع X يمكن تقريبه بالتوزيع الطبيعي مع متوسط $\mu = np$ و تباين $\sigma^2 = np(1-p)$. الشروط الملائمة هي أن تكون n كبيرة كفاية و p لا تأخذ قيمها في الأطراف (ليست قريبة من 0 أو 1). n كبيرة يضمن عدد كبير من القيم المتقطعة لمقاربتها بقيم مستمرة و p لا تأخذ قيمها في الأطراف لكي يقترب توزيع ثنائي الحد من التناظر ليقترب من التوزيع الطبيعي. مثل العديد من التقريبات أو المقاربات لا توجد قاعدة عامة مطلقة حيث يجب من أجل كل حالة التحقق للتقرير إن كان التقريب جيد كفاية للحالة المدروسة. في مقاربتنا هذه بين توزيع ثنائي الحد والتوزيع الطبيعي، القاعدة المعمول بها ليكون التقريب جيدا هي تحقق $\min(np, n(1-p)) \geq 5$.

مثال 2.3. (**التقريب الطبيعي**):

ليكن $X \sim B(25;0{,}6)$. يمكننا تقريب X بمتغير عشوائي طبيعي Y، مع $\mu = 25(0{,}6) = 15$ وانحراف معياري $\sigma = ((25)(0{,}6)(0{,}4))^{1/2} = 2{,}45$. كمثال نكتب:

$$P(X \leq 13) \approx P(Y \leq 13) = P\left(Z \leq \frac{13-15}{2{,}45}\right) = P(Z \leq -0{,}82) = 0{,}206$$

أين يكون الحساب الدقيق باستعمال دالة كثافة توزيع ثنائي الحد على النحو:

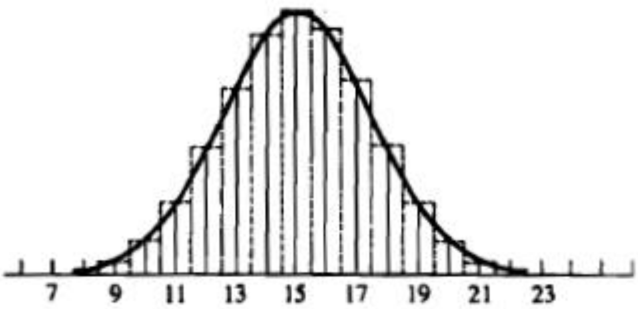
$$P(X \leq 13) = \sum_{x=0}^{13} C_{25}^{x}(0{,}6)^x(0{,}4)^{25-x} = 0{,}268$$

المقارنة تظهر أن التقريب الطبيعي جيد لكن ليس ممتازا. يمكن تحسين التقريب بشكل أفضل بالتصحيح المستمر. لرؤية كيفية عمل هذا نلاحظ في الشكل 2.3 أدناه، الذي يبين التمثيل البياني لدالتي كثافة $X \sim B(25;0{,}6)$ و $X \sim N(15;(2{,}45)^2)$. تم رسم دالة الكثافة لتوزيع ثنائي الحد باستعمال أعمدة قواعدها طولها وحدة (1) واحدة وارتفاعها مساوي لقيمة الاحتمال الموافق، في النهاية يكون مساحة كل عمود موافق لقيمة الاحتمال الموافق. عند التقريب أو المقاربة نلاحظ من الشكل أن مسحاة التوزيع الطبيعي أصغر من مساحة توزيع ثنائي الحد الموافقة للتقارب (مساحة التوزيع الطبيعي، في مثالنا، تمثل كل المساحة التي تقع على يسار الخط العمودي عند القيمة 13 بينما تلك بالنسبة لتوزيع ثنائي الحد تحوي إضافة إلى ذلك المساحة بين الخطين العمودين الموافقين للقيمتين 13 و 13,5). التصحيح المستمر

يضيف هذه المجال للتوزيع الطبيعي فبدل التقريب بحساب $P(Y \leq 13)$ نحسب $P(Y \leq 13{,}5)$ فنتحصل على قيمة مقاربة مصححة:

$$P(X \leq 13) \approx P(Y \leq 13{,}5) = P\left(Z \leq \frac{13{,}5 - 15}{2{,}45}\right) = P(Z \leq -0{,}61) = 0{,}270.$$

التي تعطي مقاربة أكثر دقة. في العموم، قيمة التقريب الطبيعي مع التصحيح المستمر هي أكبر من قيمة التقريب الطبيعي دون التصحيح المستمر.



الشكل 2.3. تقريب $N(15;(2,45)^2)$ ل $B(25;0,6)$

يمكننا أيضا القيام بالمقاربة السابقة من جهة الطرف الأعلى. إذا كان $X \sim B(n,p)$ و $Y \sim N(\mu,\sigma^2)$ يكون التقريب بينهما بصفة عامة:

$$P(X \leq x) \approx P(Y \leq y + 1/2)$$
$$P(X \geq x) \approx P(Y \geq y - 1/2)$$

ملاحظة: هذه العلاقة تبقى صالحة حتى من أجل "$x = np$".

3.ث. توزيع بيتا *Beta Distribution*:

عائلة توزيع بيتا هي معرفة على المجال $]0,1[$ بدلالة معلمتين. دالة الكثافة الاحتمالية لتوزيع بيتا تعطى ب:

$$f(x,\alpha,\beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1, \qquad \alpha > 0, \ \beta > 0 \ldots\ldots (3.16)$$

بحيث $B(\alpha,\beta)$ تشير إلى دالة بيتا المعرفة ب:

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$$

ونكتب من أجل متغير عشوائي X يتبع توزيع بيتا $X \sim B(\alpha,\beta)$.

دالة بيتا هي مرتبطة بدالة غاما بالعلاقة التالية:

$$B(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\ldots\ldots..(3.17)$$

نبرهن ذلك:

لدينا:

$$\Gamma(\alpha)\Gamma(\beta)=\left(\int_0^{\infty}t^{\alpha-1}e^{-t}dt\right).\left(\int_0^{\infty}u^{\beta-1}e^{-u}du\right)$$
$$=\int_{u=0}^{\infty}\int_{t=0}^{\infty}\left(t^{\alpha-1}u^{\beta-1}e^{-(t+u)}dtdu\right)$$

بوضع:

$$\begin{cases}t=xy\\u=y(1-x)\end{cases}$$

يصبح:

$$\begin{cases}t+u=y\\\left(\frac{t}{t+u}\right)=x\end{cases}$$

و

$$(t,u)\in[0,\infty[\Rightarrow y\in[0,\infty[\text{ و }x\in[0,1]$$

ويصبح:

$$\Gamma(\alpha)\Gamma(\beta)=\int_{y=0}^{\infty}\int_{x=0}^{1}\left((xy)^{\alpha-1}(y(1-x))^{\beta-1}e^{-y}ydxdy\right)$$
$$=\left(\int_0^{\infty}y^{\alpha+\beta-1}e^{-y}dy\right).\left(\int_0^{1}x^{\alpha-1}(1-x)^{\beta-1}dx\right)$$
$$=\Gamma(\alpha+\beta).B(\alpha,\beta)$$

ومنه نجد العلاقة (17.3).

ملاحظة: نذكر أنه عند تغيير متغير في التكامل الثنائي نستعمل المحدد الجاكوبي على النحو التالي:

$$dtdu=\begin{vmatrix}\frac{dt}{dx}&\frac{dt}{dy}\\\frac{du}{dx}&\frac{du}{dy}\end{vmatrix}dxdy=\begin{vmatrix}y&x\\-y&(1-x)\end{vmatrix}dxdy=ydxdy$$

العلاقة (17.3) هي علاقة مفيدة للتعامل مع دالة بيتا حيث تسمح لنا بالتعامل معها مع الاستفادة من خصائص دالة غاما. أكثر من ذلك، لا نتعامل مباشرة مع دالة بيتا بل نستعمل العلاقة (17.3) من أجل جميع الحسابات الضرورية.

توزيع بيتا هو من بين التوزيعات القليلة التي تعطي احتمال 1 لمجال لمجموعة تعريف منتهية بين 0 و 1. لهذا، يستعمل، توزيع بيتا، لنمذجة النسب التي تكون بالطبع معرفة بين 0 و 1.

حساب عزوم توزيع بيتا هو سهل نسبيا، الأمر الذي يرجع إلى شكل وصيغة دالة كثافتها الاحتمالية. من أجل $n > -\alpha$ لدينا:

$$E(X^n) = \frac{1}{B(\alpha,\beta)}\int_0^1 x^n x^{\alpha-1}(1-x)^{\beta-1}dx$$

$$= \frac{1}{B(\alpha,\beta)}\int_0^1 x^{(\alpha+n)-1}(1-x)^{\beta-1}dx$$

أين يمكننا أن نلاحظ أن طرف التكامل هو نواة توزيع بيتا $B(\alpha+n,\beta)$، ومنه:

$$E(X^n) = \frac{B(\alpha+n,\beta)}{B(\alpha,\beta)} = \frac{\Gamma(\alpha+n)\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+n)\Gamma(\alpha)} \dots\dots (3.18)$$

باستعمال (3.3) و (18.3) مع $n=1$ و $n=2$ نحسب المتوسط والتباين لتوزيع بيتا $B(\alpha,\beta)$ على النحوالتالي:

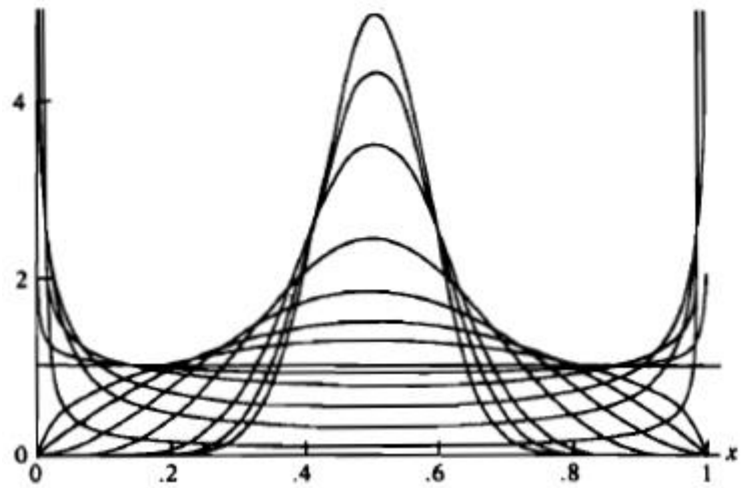
$$E(X) = \frac{\alpha}{\alpha+\beta}$$

و

$$V(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

الشكل 3.3. دوال كثافة توزيع بيتا

بما أن α و β يتغيران فإن التمثيل البياني لتوزيع بيتا يأخذ عدة أشكال مثل ما هو موضح في الشكل (3.3) أعلاه. يمكن أن تكون دالة الكثافة الاحتمالية متزايدة تماما ($\alpha > 1, \beta = 1$)، متناقصة تماما ($\alpha = 1, \beta > 1$)، على شكل حرف U ($\alpha < 1, \beta < 1$) أو وحيدة المنوال ($\alpha > 1, \beta > 1$). حالة ($\alpha = \beta$) تعطي شكل توزيع متناظر بالنسبة ل $\frac{1}{2}$ مع متوسط $\frac{1}{2}$ (بالضرورة) وتباين $\left(4(2\alpha + 1)\right)^{-1}$. كما تصبح دالة الكثافة أكثر تمركزا كلما زادت قيمة لكن تبقى متناظرة مثلما هو مبين في الشكل (4.3) أسفله. أخيرا، لما ($\alpha = \beta = 1$)، يختزل توزيع بيتا إلى التوزيع أحادي الشكل $U(0,1)$ حيث يظهر أن هذا الأخير يمكن اعتباره كحالة خاصة من عائلة توزيع بيتا. أيضا توزيع بيتا هو مرتبط، من خلال تحويل ملائم، بتوزيع فيشر (***Fisher Distribution***)، التوزيع ذو الأهمية البالغة في نظرية الاستدلال الإحصائي.



الشكل 4.3. دوال كثافة توزيع بيتا المتناظرة بالنسبة ل $(0,5)$ في حالة $(\alpha = \beta)$

3.ج. توزيع كوشي ***Cauchy Distribution***:

توزيع كوشي هو توزيع متناظر على شكل جرس معرف على مجموعة الأعداد الحقيقية $]-\infty; \infty[$ ، تعطى دالة كثافته ب:

$$f(x/\theta) = \frac{1}{\pi}.\frac{1}{1 + (x - \theta)^2}, \qquad -\infty < x < \infty, \qquad -\infty < \theta < \infty \;\ldots\ldots\ldots (3.19)$$

ونكتب من أجل متغير عشوائي X يتبع توزيع كوشي $X \sim C(\theta)$

من خلال الملاحظة بالعين المجردة (انظر على سبيل المثال الحالة الممثلة في الشكل 5.3 أسفله) يبدو أن توزيع كوشي لا يختلف عن التوزيع الطبيعي، لكن في الواقع يوجد اختلاف كبير بينهما. أولا، متوسط توزيع كوشي غير منتهي (لانهائي أو غير محدد)، كون:

$$E(|X|) = \int_{-\infty}^{\infty} \frac{1}{\pi}.\frac{|X|}{1 + (x - \theta)^2}\, dx = \infty. \ldots\ldots\ldots (3.20)$$

البرهان:

$$E(|X|) = \frac{2}{\pi}\int_0^{\infty}\frac{x}{1+(x-\theta)^2}dx$$

$$= \frac{2}{\pi}\int_0^{\infty}\frac{1}{2}.\frac{2(x-\theta)+2\theta}{1+(x-\theta)^2}dx$$

$$= \frac{2}{\pi}\left[\frac{1}{2}\int_0^{\infty}\frac{2(x-\theta)}{1+(x-\theta)^2}dx + \int_0^{\infty}\frac{\theta}{1+(x-\theta)^2}dx\right]$$

$$= \frac{2}{\pi}\left[\left[\frac{1}{2}\ln(1+(x-\theta)^2)\right]_0^{\infty} + [\theta arctg(x-\theta)]_0^{\infty}\right]$$

$$= \frac{2}{\pi}[\infty + c]$$ الدالة $arctg$ هي محدودة

$$= \infty$$

بعد هذا البرهان سيكون من السهل إثبات أن العلاقة (19.3) هي دالة كثافة احتمالية مهما تكن المعلمة θ:

$$\int_{-\infty}^{\infty} f(x/\theta)dx = \int_{-\infty}^{\infty}\frac{1}{\pi}.\frac{1}{1+(x-\theta)^2}dx = \frac{1}{\pi}[arctg((x-\theta)]_{-\infty}^{\infty} = \frac{1}{\pi}\left(\frac{\pi}{2}-\left(-\frac{\pi}{2}\right)\right)$$
$$= 1$$

بما أن $E(|X|) = \infty$ فإنه يتبع عن ذلك أنه لا توجد عزوم لتوزيع كوشي (لا توجد بالضرورة دالة توليد العزوم).

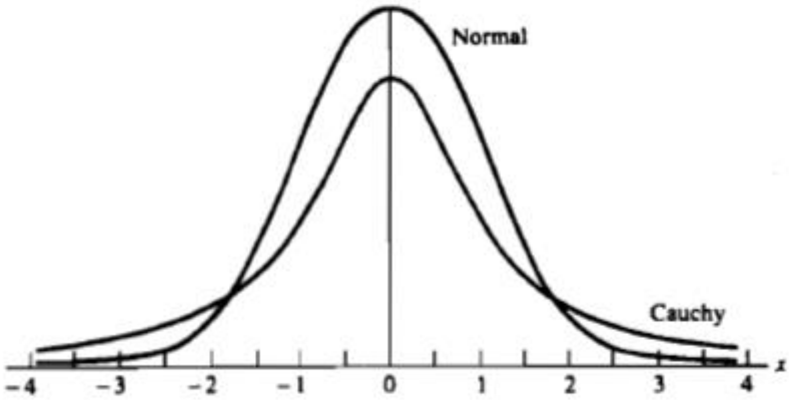
المعلمة θ في العلاقة (19.3) تعبر أو تمثل مركز التوزيع المعبر عنه بالوسيط. حيث يمكن برهنة أنه إذا كان المتغير X يخضع لتوزيع كوشي ذو معلمة θ، أي $X \sim C(\theta)$، فإنه يكون لدينا $P(X \geq \theta) = 1/2$:

$$P(X \geq \theta) = \int_{\theta}^{\infty} f(x/\theta)dx = \int_{\theta}^{\infty}\frac{1}{\pi}.\frac{1}{1+(x-\theta)^2}dx$$

$$= \frac{1}{\pi}[arctg((x-\theta)]_{\theta}^{\infty} = \frac{1}{\pi}\left(\frac{\pi}{2}-0\right) = \frac{1}{2}$$

الأمر الذي يظهر أن المعلمة θ تمثل الوسيط الحسابي للتوزيع.

الشكل (5.3) أسفله يعرض التمثيل البياني لتوزيعي كوشي $C(\theta = 0)$ و التوزيع الطبيعي المعياري $N(0,1)$ أين نرى التشابه بينهما لكن توزيع كوشي هو أكثر ثخانة في الأطراف.



الشكل 5.3. التمثيل البياني لتوزيعي كوشي $C(0)$ و التوزيع الطبيعي المعياري $N(0,1)$

3.ح.التوزيع الطبيعي اللوغاريتمي *Lognormal Distribution*:

إذا كان X متغير عشوائي حيث لوغاريتمه يخضع للتوزيع الطبيعي $lnX \sim N(\mu;\sigma^2)$ فإن X يخضع أو يتبع التوزيع الطبيعي اللوغاريتمي ونكتب $X \sim LN(\mu;\sigma^2)$. دالة كثافته الاحتمالية تعطى ب:

$$f(x/\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\frac{1}{x}e^{-(lnx-\mu)^2/(2\sigma^2)},$$
$$0 < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0 \ldots\ldots\ldots (3.21)$$

أين يمكن الحصول عليها بتحويل دالة كثافة التوزيع الطبيعي واستعمال بعض التحويلات الرياضية المناسبة على النحو التالي:

نضع $Y = lnX \sim N(\mu;\sigma^2)$ ومنه:

$$P(Y \le y) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{y} e^{-(y-\mu)^2/(2\sigma^2)}\,dy$$
$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{y} e^{-(lnx-\mu)^2/(2\sigma^2)}\,dy$$

لدينا

$$y = lnx \Rightarrow \frac{dy}{dx} = \frac{1}{x} \Rightarrow dy = \frac{dx}{x}$$

و

$$y \in]-\infty;\infty[\Rightarrow x \in]0;\infty[$$

ومنه يصبح

$$P(Y \le y) = \frac{1}{\sqrt{2\pi}\sigma}\frac{1}{x}\int_{0}^{x} e^{-(lnx-\mu)^2/(2\sigma^2)}\,dx$$

وهو المطلوب $= P(X \leq x)$

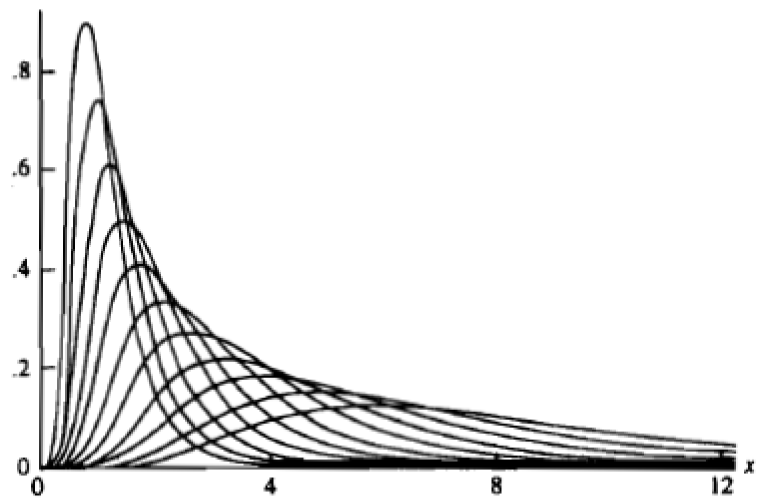
عزوم المتغير X يمكن حسابها مباشرة من العلاقة (21.3) أو باستغلال علاقته بالتوزيع الطبيعي وكتابة:

$$E(X) = E(e^{lnX})$$
$$= E(e^{Y})$$
$$= e^{\mu+(\sigma^2/2)}$$

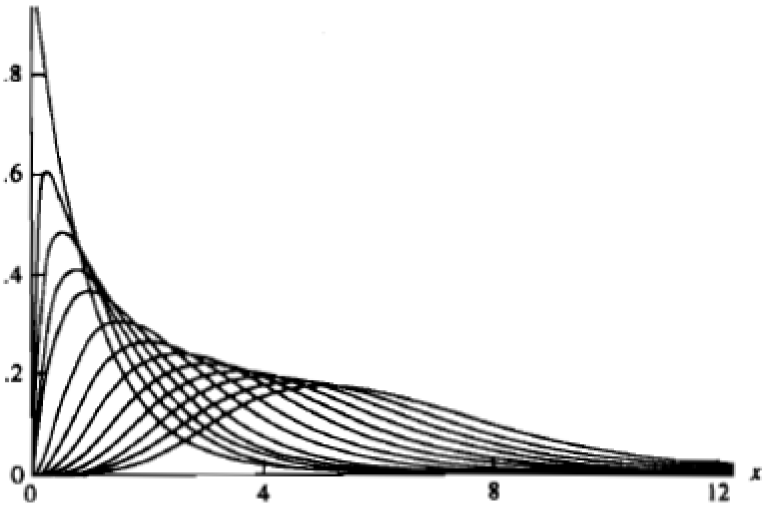
هذه الأخيرة يتحصل عليها من خلال دالة العزوم للتوزيع الطبيعي مع $t = 1$. بطريقة مماثلة نحسب $E(X^2)$ ونتحصل على:

$$V(X) = e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2}$$

التوزيع الطبيعي اللوغاريتمي يشبه في الظاهر توزيع غاما مثلما يظهره الشكل (6.3) التالي. التوزيع هو أكثر استعمالا في نمذجة التطبيقات أو المتغيرات التي يكون توزيعها ملتويا نحو اليمين. مثلا عند دراسة متغير الدخل، يكون هذا الأخير ملتويا نحو اليمين ونمذجته باستعمال التوزيع الطبيعي اللوغاريتمي تكون أكثر ملاءمة.



(أ)



(ب)

الشكل 6.3. –أ– بعض التوزيعات الطبيعية اللوغاريتمية –ب– بعض توزيعات غاما.

3.خ.التوزيع الأسي المزدوج (المضاعف) *Double Exponential Distribution*:

يصاغ التوزيع الأسي المزدوج (المضاعف) من خلال تمثيله وعكسه لسلوك التوزيع الأسي حول متوسطه. تعطى دالة كثافته الاحتمالية ب:

$$f(x/\mu,\sigma)=\frac{1}{2\sigma}e^{-|x-\mu|/\sigma},\quad -\infty<x<\infty,\quad -\infty<\mu<\infty,\quad \sigma>0 \quad \dots\dots(3.22)$$

$$\int_{-\infty}^{\infty}f(x/\mu,\sigma)dx=\int_{-\infty}^{\mu}\frac{1}{2\sigma}e^{(x-\mu)/\sigma}dx+\int_{\mu}^{\infty}\frac{1}{2\sigma}e^{-(x-\mu)/\sigma}dx$$

$$=\frac{1}{2\sigma}\left(\left[\sigma e^{(x-\mu)/\sigma}\right]_{-\infty}^{\mu}+\left[-\sigma e^{-(x-\mu)/\sigma}\right]_{\mu}^{\infty}\right)$$

$$=\frac{1}{2\sigma}(\sigma+\sigma)=1$$

نكتب بالنسبة لمتغير عشوائي X يتبع التوزيع الأسي المضاعف $X \sim DE(\mu,\sigma)$

التوزيع الأسي المزدوج (المضاعف) هو توزيع متناظر مع أطراف ثخينة (أكثر ثخانة من التوزيع الطبيعي) لكنه يبقى محتفظ بجميع عزومه عكس توزيع كوشي. يمكن حساب أمله الرياضي وتباينه للحصول على:

$$E(X)=\mu \quad \text{و} \quad V(X)=2\sigma^2$$

تمثيل التوزيع الأسي المزدوج (المضاعف) لا يأخذ شكل جرس ودالته غير قابلة للاشتقاق عند "$x = \mu$" حيث يجب تذكر هذه الخاصية عند التعامل التحليلي معه. أيضا، عند حساب التكاملات المرتبطة به من الأفضل التخلص من القيمة المطلقة، مثلما فعلنا أعلاه، من خلال تقسيم التكامل حول $x = \mu$:

$$E(X) = \int_{-\infty}^{\infty} \frac{x}{2\sigma} e^{-|x-\mu|/\sigma} dx$$

$$= \int_{-\infty}^{\mu} \frac{x}{2\sigma} e^{(x-\mu)/\sigma} dx + \int_{\mu}^{\infty} \frac{x}{2\sigma} e^{-(x-\mu)/\sigma} dx \ldots\ldots (3.23)$$

حساب تكاملات العلاقة (23.3) يتم بالتكامل بالتجزئة لكل جزء.

نذكر في النهاية إلى أنه توجد عدة توزيعات مستمرة أخرى تستعمل في مختلف التطبيقات الإحصائية. يمكن الرجوع إلى الكتاب القيم لجونسون و كوتز "*Johnson and Kotz (1969-1972)*" "*Distributions in Statistics*" من أجل التعرف على توزيعات إحصائية أخرى.

4. العائلات الأسية *Exponential Families*:

يقال عن فئة من دوال الكثافة الاحتمالية، $f(x/\theta)$، أنها فئة أسية إذ كان يمكن التعبير عنها ب:

$$f(x/\theta) = h(x)c(\theta)exp\vartheta\left(\sum_{i=1}^{k} w_i(\theta)\, t_i(x)\right) \ldots\ldots.(4.1)$$

بحيث $h(x) > 0$ و $t_i(x), i = \overline{1,k}$ هي دوال حقيقية معرفة بالمشاهدة x (لا يمكن ان تكون تابعة أو مرتبطة ب θ)، $c(\theta) > 0$ و $w_i(\theta), i = \overline{1,k}$ هي دوال حقيقية معرفة بدلالة شعاع المعالم θ (لا يمكن ان تكون تابعة أو مرتبطة ب x). في الواقع، العديد من التوزيعات المتطرق إليها سابقا هي من الفئة الأسية. هذا يشمل التوزيعات المتقطعة –ثنائي الحد، بواسون و ثنائي الحد السالب– و التوزيعات المستمرة– الطبيعي، غاما وبيتا–.

للتحقق من دالة توزيع أنها عائلة أسية يجب تحديد الدوال $h(x)$، $c(\theta)$، $w_i(\theta)$ و $t_i(x)$ من أجل $i = \overline{1,k}$ وإظهار أنه يمكن كتابتها من الشكل (1.4). المثال التالي يوضح ذلك.

مثال 1.4. (**عائلـة ثنائي الحد الأسية**):

ليكن n عدد طبيعي و $B(n,p)$ عائلة ثنائي الحد مع $0 < p < 1$. إذا دالة الكثافة الاحتمالية لهذه العائلة، من أجل $x = 0,\ldots,n$ و $0 < p < 1$، يعبر عنها على النحو التالي:

$$\begin{aligned} f(x/p) &= \binom{x}{n} p^x (1-p)^{n-x} \\ &= \binom{x}{n} (1-p)^n \left(\frac{p}{1-p}\right)^x \\ &= \binom{x}{n} (1-p)^n exp\left(ln\left(\frac{p}{1-p}\right).x\right) \end{aligned}$$

نعرف

$$h(x) = \begin{cases} \binom{x}{n} & x = 0,\ldots,n \\ 0 & \text{عدا ذلك} \end{cases} c(p) = (1-p)^n, \qquad 0 < p < 1$$

$$w_1(p) = ln\left(\frac{p}{1-p}\right), \qquad 0 < p < 1 \;\text{و}\; \quad t_1(x) = x$$

أين يكون لدينا في النهاية

$$f(x/p) = h(x)c(p)exp\big(w_1(p)t_1(x)\big) \ldots\ldots\ldots(4.2)$$

التي تأخذ شكل العلاقة (1.4) مع $k = 1$. نشير بالخصوص أنه يكون "$h(x) > 0$" فقط لما "$x = 0, \dots, n$" و تكون $c(p)$ معرفة فقط لما يكون "$0 < p < 1$" . هذا مهم حتى تتوافق (3.4) مع (2.4) من أجل جميع قيم x وتكون عائلة أسية فقط لما "$0 < p < 1$" (إذا دوال المعلمة هي معرفة فقط هنا). نضيف، قيمتي المعلمة $p = 0$ و $p = 1$ أحيانا يتم دمجهما في نموذج ثنائي الحد، لكن لم يتم إدماجهما هنا لأن مجموعة قيم x التي من أجلها تكون "$f(x/p) > 0$" هي مختلفة من أجل 0 و $p = 1$ عن تلك القيم مع باقي قيم p.

الصيغة الخاصة (1.4) تعطي العائلة الأسية خصائص رياضية مفيدة. لكن أكثر أهمية بالنسبة للنماذج الإحصائية، الصيغة (1.4) تعطي العائلة الأسية خصائص إحصائية رائعة. نوضح فيما يلي حسابات مختصرة لعزومها.

نظرية 1.4.:

ليكن X متغير عشوائي له دالة كثافة احتمالية من الشكل (1.4)، إذا يكون لدينا:

$$E\left(\sum_{i=1}^{k}\frac{\partial w_i(\theta)}{\partial\theta_j}t_i(X)\right) = -\frac{\partial}{\partial\theta_j}lnc(\theta); \dots\dots\dots (4.3)$$

$$V\left(\sum_{i=1}^{k}\frac{\partial w_i(\theta)}{\partial\theta_j}t_i(X)\right) = -\frac{\partial^2}{\partial\theta_j^2}ln\, c(\theta) - E\left(\sum_{i=1}^{k}\frac{\partial^2 w_i(\theta)}{\partial\theta_j^2}t_i(X)\right); \dots\dots\dots (4.4)$$

ميزة هذه العلاقات أنه من خلالها يمكن استبدال التكاملات أو المجاميع بالمشتقات، الأمر الذي يكون في الغالب أكثر سهولة.

مثال 2.4. (**متوسط وتباين توزيع ثنائي الحد**):

من المثال السابق لدينا:

$$\frac{d}{dp}w_1(p) = \frac{d}{dp}ln\frac{p}{1-p} = \frac{1}{p(1-p)}$$

$$\frac{d}{dp}lnc(p) = \frac{d}{dp}nln(1-p) = \frac{-n}{1-p}$$

ومنه حسب النظرية:

$$E\left(\frac{1}{p(1-p)}X\right) = \frac{n}{1-p}$$

مع بعض الترتيبات نجد $E(X) = np$. صيغة التباين يتحصل عليها بطريقة مماثلة.

نتناول الآن مثال آخر ومجموعة من العائلات الأسية المهمة.

مثال 3.4. (**العائلة الأسية الطبيعية**):

لتكن $f(x/\mu,\sigma^2)$ دوال الكثافة الاحتمالية لعائلة التوزيع الطبيعي $N(\mu,\sigma^2)$ أين $\theta=(\mu,\sigma^2)$، $-\infty<\mu<\infty$، $\sigma>0$. إذا:

$$f(x/\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma}exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$
$$=\frac{1}{\sqrt{2\pi}\sigma}exp\left(\frac{-\mu^2}{2\sigma^2}\right)exp\left(\frac{-x^2}{2\sigma^2}+\frac{\mu x}{\sigma^2}\right)\dots\dots(4.5)$$

نعرف

$h=1,$ من أجل كل x

$$c(\theta)=c(\mu,\sigma)=\frac{1}{\sqrt{2\pi}\sigma}exp\left(\frac{-\mu^2}{2\sigma^2}\right),-\infty<\mu<\infty,\qquad \sigma>0;$$

$$w_1(\theta)=w_1(\mu,\sigma)=\frac{1}{\sigma^2},\qquad \sigma>0;\;\; w_2(\theta)=w_2(\mu,\sigma)=\frac{\mu}{\sigma^2}\;,\qquad \sigma>0;$$

$t_1(x)=-x^2/2$ و $t_2(x)=x$

إذا

$$f(x/\mu,\sigma^2)=h(x)c(\mu,\sigma)exp\big(w_1(\mu,\sigma)t_1(x)+w_2(\mu,\sigma)t_2(x)\big)$$

التي هي من الشكل (1.4) مع $k=2$. نشير مجددا أن دوال المعالم هي معرفة فقط بدلالة مجموعة المعالم.

في العموم، مجموعة قيم x في العائلة الأسية ، التي من أجلها $f(x/\theta)>0$، لا يمكن أن تكون مرتبطة ب θ. كل عناصر أو تعريف دالة الكثافة الاحتمالية يجب أن يدمج في عناصر الصيغة (1.4). يكون هذا أسهل من خلال دمج عناصر x في صيغة $f(x/\theta)$ من خلال استعمال دالة التأشير (الدالة المؤشرة).

تعريف 1.4:

دالة التأشير لمجموعة A ، التي يرمز لها غالبا ب $I_A(x)$، هي الدالة:

$$I_A(x)=\begin{cases}x\;\in A\\ x\notin A\end{cases}$$

ترميز بديل لدالة التأشير $I(x\;\in A)$.

دالة الكثافة الاحتمالية للتوزيع الطبيعي للمثال السابق يمكن أن تكتب:

$$f(x/\mu,\sigma^2)=h(x)c(\mu,\sigma)exp\big(w_1(\mu,\sigma)t_1(x)+w_2(\mu,\sigma)t_2(x)\big)I_{]-\infty;\infty[}(x)$$

بما أن دالة التأشير هي دالة تابعة فقط ل x ، فيمكن إذا دمجها مع الدالة $h(x)$ حيث تصبح هذه الكتابة الأخيرة لدالة الكثافة الاحتمالية هي أيضا من الشكل (1.4).

من خلال (1.4)، كون المعامل $\exp(.)$ موجب دائما، يمكن ملاحظة أنه من أجل كل $\theta \in \Theta$، ومن أجل أي θ التي من اجلها $c(\theta) > 0$، $\{x: f(x/\theta) > 0\} = \{x: h(x) > 0$ وهذه المجموعة هي غير مرتبطة ب θ . مثلا، دوال الكثافة الاحتمالية المعطاة بالعلاقة:

$$f(x/\theta) = \theta^{-1} \exp\big(1 - (x/\theta)\big), 0 < \theta < x < \infty$$

هي ليست عائلة أسية على الرغم من إمكانية كتابتها من الشكل:

$h(x)c(\theta)exp\big(w_1(\theta)t_1(x)\big)$ مع $h(x) = e^1$، $c(\theta) = \theta^{-1}$، $w_1(\theta) = \theta^{-1}$ و $t_1(x) = -x$.

كتابة دالة الكثافة الاحتمالية بدلالة دالة التأشير يجعل ذلك واضحا ومبررا. لدينا:

$$f(x/\theta) = \theta^{-1} \exp\big(1 - (x/\theta)\big)\, I_{]\theta;\infty[}(x)$$

دالة التأشير لا يمكن دمجها في أي عنصر من (1.4) كونها ليست تابعة لا ل x لوحده ولا ل θ وحدها ولا يمكن كتابتها على شكل أسي. الأمر الذي يجعلها ليست عائلة أسية.

تكتب العائلة الأسية على الشكل:

$$f(x/\eta) = h(x)c^*(\eta)exp\left(\sum_{i=1}^{k} \eta_i t_i(x)\right) \ldots\ldots (4.6)$$

هنا، الدوال $h(x)$ و $t_i(x)$ هي نفسها المعطاة في العلاقة (1.4). المجموعة $\mathcal{H} = \{\eta = (\eta_i, \ldots\ldots, \eta_k): \int_{-\infty}^{\infty} h(x)xp\big(\sum_{i=1}^{k} \eta_i t_i(x)\big)dx < \infty\}$ تسمى **بالفضاء الطبيعي للمعالم** للعائلة. (إذا كان X متغير متقطع، يستبدل التكامل بالمجموع من أجل قيم x التي من أجلها $h(x) > 0$). من أجل قيم $\eta \in \mathcal{H}$ يجب أن يكون لدينا $c^*(\eta) = \left[\int_{-\infty}^{\infty} h(x)xp\big(\sum_{i=1}^{k} \eta_i t_i(x)\big)dx\right]^{-1}$ لضمان أن يكون تكامل دالة الكثافة الاحتمالية مساوي ل 1. كون الدالة الأصلية $f(x/\theta)$ في (1.4) هي دالة كثافة احتمالية، المجموعة $\{\eta = \big(w_1(\theta), \ldots, w_k(\theta)\big): \theta \in \Theta\}$ يجب ان تكون مجموعة جزئية من فضاء المعالم. لكن توجد قيم أخرى ل $\eta \in \mathcal{H}$.

مثال 4.4. (تابع للمثال السابق):

لتحديد فضاء المعالم الطبيعية لعائلة التوزيع الطبيعي، بتعويض $w_i(\mu, \sigma)$ ب η_i في (6.4) نتحصل على:

$$f(x/\eta_1, \eta_2) = \frac{\sqrt{\eta_1}}{\sqrt{2\pi}} exp\left(-\frac{\eta_2^2}{2\eta_1}\right) exp\left(-\frac{\eta_1 x^2}{2} + \eta_2 x\right) \ldots\ldots (4.7)$$

يكون التكامل منتهي إذا وفقط إذا كان معامل x^2 سالب. هذا يعني أن يكون η_1 موجب. إذا كان $\eta_1 > 0$ يكون التكامل منتهي بغض النظر عن قيمة η_2. بهذا يكون فضاء المعالم الطبيعية هو $\{(\eta_1, \eta_2): \eta_1 > 0, -\infty < \eta_2 < \infty\}$. بمطابقة (7.4) مع

(5.4)، نلاحظ ان $\eta_2 = \mu/\sigma^2$ و $\eta_1 = 1/\sigma^2$. على الرغم من المزايا الرياضية للمعالم الطبيعية، تكون غير كافية لبعض التفسيرات البسيطة مثل المتوسط والتباين.

في الصيغة (1.4) يكون غالبا بعد شعاع المعالم θ يساوي k، عدد العناصر في المجموع الأسي. لكن ليس بالضرورة حيث يمكن أن يكون بعد شعاع المعالم θ يساوي d و $d < k$. هذا الصنف من العائلات الأسية يسمى **العائلة الأسية المنحنية**.

تعريف 2.4:

العائلة الأسية المنحنية هي عائلة كثافتها من الشكل (1.4) ويكون بعد شعاع المعالم θ يساوي d و $d < k$. إذا كان $d = k$ فهي عائلة أسية تامة.

مثال 5.4. (**عائلة أسية منحنية**):

العائلة الطبيعية للمثال السابق هي عائلة أسية تامة. فرضا لو نضع $\sigma^2 = \mu^2$، تصبح منحنية. (النماذج المماثلة ممكن أن تستعمل في تحليل التباين) يكون لدينا إذا:

$$f(x\mu) = \frac{1}{\sqrt{2\pi\mu^2}} exp\left(\frac{-(x-\mu)^2}{2\mu^2}\right)$$

$$= \frac{1}{\sqrt{2\pi\mu^2}} exp\left(\frac{-1}{2}\right) exp\left(\frac{-x^2}{2\mu^2} + \frac{x}{\mu}\right) \dots \dots (4.8)$$

من أجل العائلة الطبيعية، العائلة الأسية التامة يجب أن يكون لها فضاء معالم $(\mu, \sigma^2) = \Re \times]0;\infty[$، بينما فضاء المعالم للعائلة المنحنية $(\mu, \sigma^2) = (\mu, \mu^2)$ هي قطع مكافئ.

العائلات الأسية المنحنية لها استعمالات عديدة. المثال التالي يوضح أحد استعمالاتها البسيطة.

مثال 6.4. (**التقريبات الطبيعية**):

إذا كان $X_1, \dots, X_n$ يمثل عينة من مجتمع بواسوني $P(\lambda)$، إذا التوزيع $\bar{X} = \sum_i X_i/n$ هو بالتقريب

$$\bar{X} \backsim N(\lambda, \lambda/n)$$

عائلة أسية منحنية.

التقريب $N(\lambda, \lambda/n)$ هو مبرر بنظرية النهايات المركزية. نشير إلى أنه أغلب تقريبات نظرية النهايات المركزية ينتج عنها عائلة طبيعية منحنية. رأينا التقريب الطبيعي لثنائي الحد (مثال (4.2.)): إذا كانت $X_1, \dots, X_n$ متغيرات مستقلة وتخضع لتوزيع برنولي $B(p)$، فإن:

$$\bar{X} \backsim N(p, p(1-p)/n)$$

هو تقريب طبيعي منحني (من أجل تفصيل أكثر حول العائلات الأسية، انظر (*Lehmane 1986*)، (*lehmann and Casella 1998*)، (*Brown 1986*)).

5. عائلات الموضع والسلم:

في العنصرين 3 و 4 تطرقنا إلى العديد العائلات من التوزيعات المستمرة. في هذا العنصر سوف نناقش 3 تقنيات لبناء مجموعة أو عائلة من التوزيعات. العائلات الناتجة بالضرورة لها تفسير تركيبي الأمر الذي يجعل منها مفيدة لنمذجة الخصائص الرياضية الملائمة.

الأصناف الثلاث تسمى عائلات الموضع (التموضع)، عائلات السلم (القياس–المدى) و عائلات الموضع–السلم. كل عائلة يتم بناؤها من خلال تشخيص دالة كثافة احتمالية بسيطة $f(x)$، تسمى بدالة الكثافة الاحتمالية العادية للعائلة. ثم كل دوال الكثافة الاحتمالية الأخرى للعائلة يتحص عليها بتحويلها إلى شكل ملائم. نبدأ بنظرية بسيطة حول دوال الكثافة الاحتمالية.

نظرية 1.5. :

لتكن $f(x)$ أي دالة كثافة احتمالية ، $0 < \mu$ و $0 < \sigma$ أي ثابتين معطين. إذا الدالة:

$$g(x,\mu,\sigma)=\frac{1}{\sigma}\,f\left(\frac{x-\mu}{\sigma}\right)$$

هي دالة كثافة احتمالية.

البرهان:

للتأكد من أن التحويل السابق أنتج دالة كثافة احتمالية، نحتاج أن نبين أن $(1/\sigma)\ f((x-\mu)/\sigma)$، كدالة ل x، هي دالة كثافة احتمالية من أجل جميع قيم μ و σ الممكن أن تأخذهم الدالة. هذا يعني تبيين أن $(1/\sigma)\ f((x-\mu)/\sigma)$ هي غير سالبة وتكاملها بالنسبة لمجموعة تعريفها يساوي 1. بداية، كون $f(x)$ هي دالة كثافة احتمالية فإن $f(x)\geq 0$ من اجل جميع قيم x. ومنه يكون أيضا، $\big((1/\sigma)\ f((x-\mu)/\sigma)\geq 0\big)$ من أجل جميع قيم x، μ و σ. أيضا:

$$\int_{-\infty}^{\infty}\frac{1}{\sigma}\,f\left(\frac{x-\mu}{\sigma}\right)dx=\int_{-\infty}^{\infty}\frac{1}{\sigma}\,f(y)\,dx,\qquad \text{بوضع}\ \ y=\frac{x-\mu}{\sigma}$$

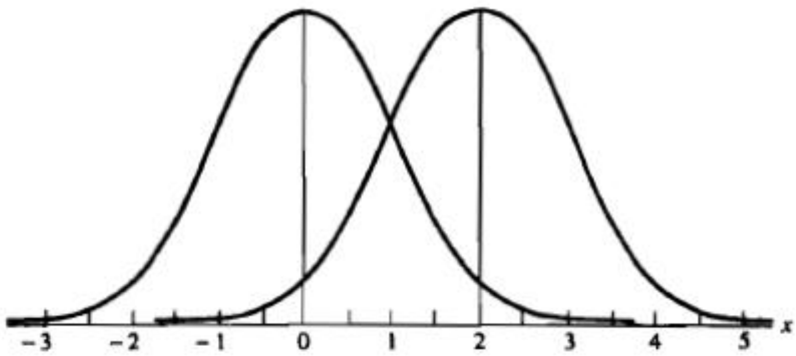
$$=\int_{-\infty}^{\infty}\frac{1}{\sigma}\,f(y)\sigma\,dy$$

$$=\int_{-\infty}^{\infty}f(y)\,dy=1,\qquad (f(y)\ \text{هي دالة كثافة احتمالية})$$

وهو المطلوب.

تعريف 1.5.:

لتكن $f(x)$ أي دالة كثافة احتمالية. إذا عائلة (مجموعة صنف) دوال الكثافة الاحتمالية $f(x-\mu)$، المؤشرة بالمعلمة μ، $-\infty<\mu<\infty$ تسمى **بعائلات الموضع (التموضع)** ذات دالة كثافة احتمالية عادية (بسيطة) $f(x)$ و μ يسمى **بمعلمة الموضع (التموضع)** للعائلة.



الشكل 1.5. توزيعين لنفس عائلة الموضع مع متوسطين 0 و 1.

لمشاهدة تأثير إدخال معلمة الموضع μ، نلاحظ الشكل (1.5) أعلاه. عند $x = \mu$ ، $f(x-\mu) = f(0)$، عند $x = \mu + 1$، $f(x-\mu) = f(1)$، وبصفة عامة، عند $x = \mu + a$، $f(x-\mu) = f(a)$. بالطبع، من أجل $\mu = 0$، $f(x-\mu)$ هي نفسها $f(x)$. في الواقع معلمة الموضع μ ببساطة تحول أو تسحب دالة التوزيع التراكمية $f(x)$ حيث شكل المنحنى البياني لا يتغير لكن النقطة من المنحنى التي تقع فوق $x = 0$ من أجل$f(x)$ نقع فوق $x = \mu$ من أجل $f(x-\mu)$. واضح من الشكل (1.5) أن المساحة تحت منحنى $f(x)$ بين $x = -1$ و $x = 2$ هي نفسها المساحة تحت منحنى $f(x-\mu)$ بين $x = \mu - 1$ و $x = \mu + 2$.

إذا كان X متغير عشوائي ودالة كثافته الاحتمالية $f(x-\mu)$، يمكننا كتابة:

$$P(-1 \leq X \leq 2/0) = (\mu - 1 \leq X \leq \mu + 2/\mu),$$

أين المتغير العشوائي X له دالة الكثافة الاحتمالية $f(x-0) = f(x)$ الممثلة على يسار المساواة ودالة الكثافة الاحتمالية $f(x-\mu)$ الممثلة على اليمين.

العديد من العائلات المذكورة في الجزء 4 لها عائلات موضع أو جزء من عائلات موضع. مثلا، من أجل العدد المحدد والمعلوم $\sigma > 0$ نعرف:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/(2\sigma^2)}, \quad -\infty < x < \infty,$$

إذا عائلة الموضع لدالة الكثافة الاحتمالية هذه هي مجوعة التوزيعات الطبيعية بمتوسط مجهول μ وتباين معلوم σ^2. لملاحظة ذلك يكفي استبدال x ب $x - \mu$ في الدالة السابقة للحصول على تلك المعطاة ب (13.3). كذلك، عائلة كوشي والعائلة الأسية المضاعفة، مع σ محددة القيمة و المعلمة μ، هي أمثلة لعائلات الموضع. التعريف (1.5.) يعني أنه يمكننا البدء بأي دالة كثافة احتمالية $f(x)$ وتوليد عائلات موضع من خلال إدخال معلمة موضع عليها.

إذا كان X متغير عشوائي دالة كثافته الاحتمالية $f(x-\mu)$، إذا يمكن تمثيل X ب $X = Z + \mu$ أين Z هو متغير عشوائي دالة كثافته الاحتمالية هي $f(Z)$.. أهمية هذا التمثيل يظهر عندما تكون عائلة موضع أكثر ملاءمة لنمذجة حالة متغير عشوائي مشاهد X.

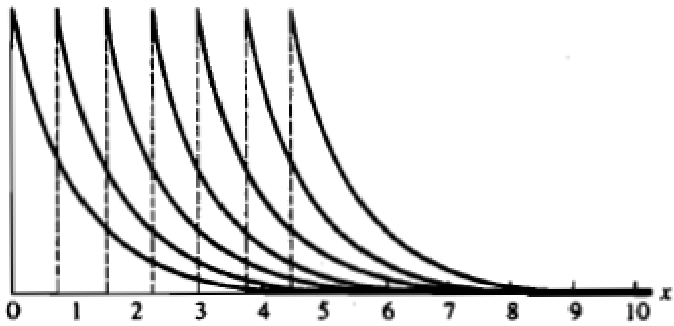
مثال 1.5. (**عائلة الموضع الأسية**):

ليكن $f(x) = e^{-x}, x \geq 0$ و $f(x) = 0$ لما $x < 0$. لصياغة أو تشكيل عائلة موضع نعوض x ب $x - \mu$ ونتحصل على:

$$f(x/\mu) = \begin{cases} e^{-(x-\mu)}, & x - \mu \geq 0 \\ 0, & x - \mu < 0 \end{cases}$$

$$= \begin{cases} e^{-(x-\mu)}, & x \geq \mu \\ 0, & x < \mu \end{cases}$$

منحنيات $f(x/\mu)$ من أجل قيم مختلفة هي ممثلة في الشكل (2.5) أسفله. مثل الشكل (1.5) السابق، التمثيل أو المنحنى تم سحبه حيث الجزء الموجب يبدأ من μ يدل من عند 0. إذا كان المتغير X يقيس الوقت أو الزمن، فإن μ يقيد بأن لا يكون سالبا حيث يكون X موجب مع احتمال كلي يساوي 1 من أجل كل قيمة ل μ. في هذا النوع من النماذج، أين μ يشير أو يمثل حد لمجموعة X، يسمى μ أحيانا **بالمعلمة العتبة**.



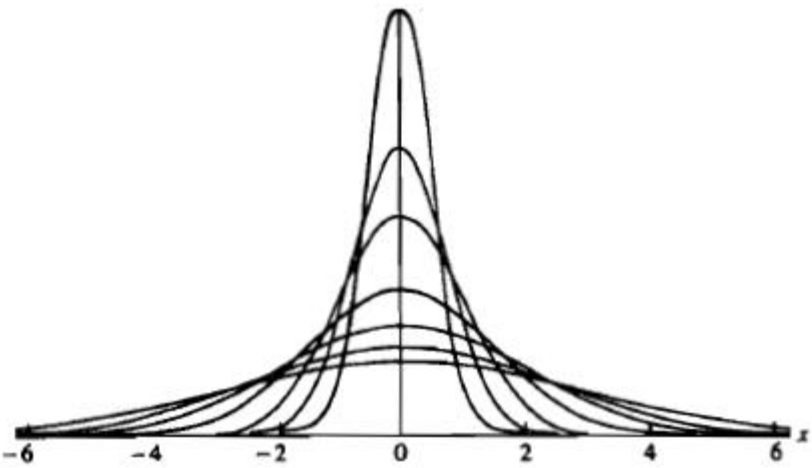
الشكل 2.5. كثافة عائلة الموضع الأسية

الصنفين الآخرين، عائلات السلم (القياس-المدى) و عائلات الموضع-السلم، سنتطرق إليهما في الفقرات التالية.

تعريف 2.5.:

لتكن $f(x)$ أي دالة كثافة احتمالية. من أجل أي $\sigma > 0$، عائلة دوال الكثافة الاحتمالية من الشكل $(1/\sigma)f(x/\sigma)$، المؤشرة بالمعلمة σ ، تسمى **بعائلات السلم** (**القياس–المدى**) ذات دالة الكثافة الاحتمالية الأساسية $f(x)$ و تسمى المعلمة σ **معلمة السلم** (**القياس، المدى**) للعائلة.

يتجلى تأثير إدخال معلمة القياس σ إما بتمديد $(\sigma > 1)$ أو بتقليص $(\sigma < 1)$ منحنى $f(x)$ مع البقاء النسبي لنفس الشكل العام للمنحنى. الشكل (3.5) التالي يوضح ذلك. تستعمل معالم السلم (القياس) غالبا عندما تكون الدالة الأساسية $f(x)$ إما متناظرة حول "0" أو موجبة فقط من أجل $x > 0$. في هذه الحالة يكون التمدد إما تناظريا بالنسبة ل "0" أو فقط في الاتجاه الموجب. لكن تعريفيا، أي دالة كثافة احتمالية يمكن أن تكون أساسية. منها عائلة غاما لما تكون قيمة α ثابتة و β معلمة سلم (قياس) و العائلة الطبيعية لما $\mu = 0$ و σ معلمة سلم (قياس). في كل حالة من هذه دالة الكثافة الاحتمالية الأساسية يتحصل عليها بوضع $\sigma = 1$. كل العائلات الأخرى يمكن أن تبين أنها من شكل التعريف (2.5).

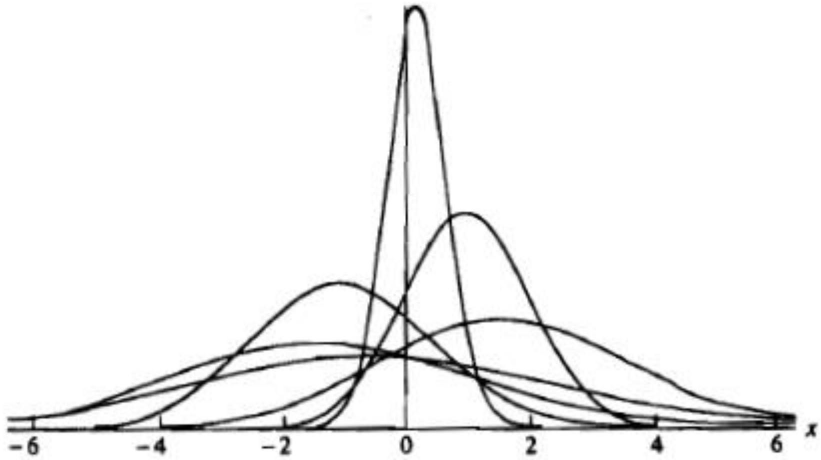


الشكل 3.5. دوال لنفس عائلة الموضع.

تعريف 3.5.:

لتكن $f(x)$ أي دالة كثافة احتمالية. إذا من أجل أي μ، $-\infty < \mu < \infty$، وأي $\sigma > 0$، عائلة دوال الكثافة الاحتمالية $(1/\sigma)f((x-\mu)/\sigma)$، المؤشرة بالمعلمتين (μ,σ)، تسمى **عائلة الموضع–السلم (القياس)** ذات دالة الكثافة الأساسية $f(x)$، حيث يسمى μ **بمعلمة الموضع** و σ **بمعلمة السلم (القياس)**.

تأثير إدخال معلمتي الموضع والقياس يكون بالتمديد $(\sigma > 1)$ أو بالتقليص $(\sigma < 1)$ منحنى $f(x)$ بالنسبة لمعلمة السلم (القياس) ويكون بسحبه بالنسبة لمعلمة الموضع. الشكل (4.5) يوضح هذه التأثيرات على منحنى $f(x)$. عائلتي الاسية المضاعف (المزدوج) و الطبيعية هي أمثلة لعائلة الموضع–السلم (القياس).



الشكل 4.5. دوال لنفس عائلة الموضع–السلم (القياس).

النظرية التالية تربط تحويل دالة الكثافة الاحتمالية التي تعرف عائلة الموضع–السلم (القياس) بتحويل متغير عشوائي Z ذو دالة كثافة احتمالية $f(z)$. مثلما تم الإشارة إليه خلال عرض عائلات الموضع، التمثيل بدلالة Z هو طريقة أو آداة رياضية مفيدة ويمكن أن تساعدنا

على فهم متى يمكن أن تكون عائلة الموضع-السلم (القياس) أكثر ملاءمة في نمذجة سياق معين. بوضع، في النظرية (1.5)، $(\sigma = 1)$ تسقط نتائج النظرية على عائلة الموضع، وبوضع $(\mu = 0)$ تسقط نتائج النظرية على عائلة السلم (القياس).

نظرية 2.5.:

لتكن $f(.)$ إي دالة كثافة احتمالية, ليكن μ أي عدد حقيقي و σ أي عدد حقيقي موجب. إذا يكون X متغير عشوائي ذو دالة كثافة احتمالية $(1/\sigma)f((x-\mu)/\sigma)$ **إذا وفقط إذا** كان يوجد متغير عشوائي Z ذو دالة كثافة احتمالية $f(z)$ و $X = \sigma Z + \mu$.

البرهان:

نبرهن أولا الجزء "**إذا**"، نعرف $g(z) = \sigma Z + \mu$. إذا $X = g(z)$، g هي دالة وحيدة الاتجاه (متزايدة أو متناقصة تماما)، إذا $g^{-1}(x) = (x-\mu)/\sigma$ و $|(d/dx)g^{-1}(x)| = 1/\sigma$، ومنه تعطى دالة الكثافة الاحتمالية ل X ب:

$$f_X(x) = f_Z\big(g^{-1}(x)\big)\left|\left(\frac{d}{dx}\right)g^{-1}(x)\right| = f\left(\frac{x-\mu}{\sigma}\right)\frac{1}{\sigma}.$$

نبرهن ثانيا الجزء "**وفقط إذا**". نعرف $g(x) = (x-\mu)/\sigma$ وليكن $Z = g(X)$. لدينا $g^{-1}(z) = \sigma Z + \mu$، $|(d/dz)g^{-1}(z)| = \sigma$ ودالة الكثافة الاحتمالية ل Z هي:

$$f_Z(z) = f_X\big(g^{-1}(z)\big)\left|\left(\frac{d}{dz}\right)g^{-1}(z)\right| = \frac{1}{\sigma}f\left(\frac{(\sigma Z+\mu)-\mu}{\sigma}\right)\sigma = f(z).$$

أيضا

$$\sigma Z + \mu = \sigma g(X) + \mu = \sigma\left(\frac{X-\mu}{\sigma}\right) + \mu = X$$

نتيجة مهمة يمكن استخلاصها من النظرية (6.5) هي أن المتغير العشوائي $Z = (X-\mu)/\sigma$ له دالة كثافة احتمالية

$$f_Z(z) = \frac{1}{1}f\left(\frac{z-0}{1}\right)\sigma = f(z).$$

هذا، توزيع Z هو عنصر من عائلة الموضع-السلم (القياس) الموافق ل $\mu = 0$ و $\sigma = 1$.

مثل ما أشرنا إليه سابقا، غالبا ما تتم الحسابات بدالة الكثافة الاحتمالية للمتغير العشوائي المعياري Z وتستنتج بسهولة النتيجة الموافقة للمتغير X ذو دالة الكثافة الاحتمالية $(1/\sigma)f((x-\mu)/\sigma)$.

نظرية 3.5.:

ليكن Z متغير عشوائي مع دالة كثافة احتمالية $f(z)$. نفرض بوجود $E(X)$ و $V(X)$. إذا كان X متغير عشوائي ذو دالة كثافة احتمالية $(1/\sigma)f((x-\mu)/\sigma)$، فإن:

$$E(X) = \sigma E(Z) + \mu \text{ و } V(X) = \sigma^2 V(Z).$$

ويكون خصوصا لما $E(Z) = 0$ و $V(Z) = 1$، $E(X) = \mu$ و $V(X) = \sigma^2$.

البرهان:

من خلال النظرية (6.5)، يوجد متغير عشوائي Z^* ذو دالة كثافة احتمالية $f(z)$ و $X = \sigma Z^* + \mu$. إذا $E(X) = \sigma E(Z^*) + \mu$ و $V(X) = \sigma^2 V(Z^*)$.

من أجل أي عائلة موضع–سلم (قياس) ذات متوسط وتباين منتهيين، دالة الكثافة الاحتمالية المعيارية $f(z)$ يمكن اختيارها بطريقة يكون $E(Z) = 0$ و $V(Z) = 1$. هذه النتيجة هي التفسير الملائم ل μ و σ^2 على انهما المتوسط والتباين على التوالي. احتمالات أي عنصر من عائلة الموضع–السلم (القياس) يمكن حسابها بدلالة أو من خلال المتغير المعياري Z لأن:

$$P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P\left(Z \leq \frac{x-\mu}{\sigma}\right).$$

لهذا، إذا تم جدولة $P(Z \leq z)$ أو تم حسابها بسهولة بالنسبة للمتغير Z، فإن الاحتمالات المتعلقة ب X يمكن إيجادها أيضا بسهولة. حساب الاحتمالات الطبيعية باستعمال الجدول الطبيعي المعياري هو مثال لهذا.

6. علاقات مساواة وثوابت:

النظرية الإحصائية هي مليئة وغنية بعلاقات المساواة والثوابت. أغلب عمل *Marshall* و *Olkin* (1979) يحوي علاقاة مساواة باستعمال مفهوم التسقيف (*majoration*). قبله، عمل *Hardy*، *Littlewood* و *Polya* (1952) يجمع علاقات المساواة الكلاسيكية. في هذه النقطة والتي تليها، سوف نجمع بين تلك علاقات المساواة الكلاسيكية و الجديدة (الحديثة)، بإعطاء بعض الأفكار حول النتائج المتوصل إليها. نشير إلى أن هذه الفقرة (أ) هي مخصصة لتلك علاقات المساواة والثوابت النابعة من نظرية الاحتمالات بينما تلك المذكورة في الفقرة الموالية (ب) هي تلك المتعلقة أكثر بخصائص الأعداد والدوال الرياضية.

6.أ. علاقات المساواة الاحتمالية:

أشهر علاقة مساواة، وربما الأكثر فائدة، هي مساواة تشيبيتشيف *Chebychev's Inequality*. فائدتها تأتي من واقع تطبيقاتها الواسعة. مثل أغلب النتائج، برهانها يكون بديهيا.

- **نظرية** (**مساواة تشيبيتشيف *Chebychev's Inequality***):

ليكن X متغير عشوائي دالة كثافته الاحتمالية $f_X(x)$ و لتكن $g(x)$ دالة غير سالبة. إذا، من أجل أي $(r > 0)$ يكون لدينا:

$$P(g(X) \geq r) \leq \frac{E\big(g(X)\big)}{r} \ldots\ldots (6.1)$$

البرهان:

$$\begin{aligned} E\big(g(X)\big) &= \int_{-\infty}^{\infty} g(X) f_X(x)dx \\ &\geq \int_{\{x:g(x)\geq r\}}^{\infty} g(X) f_X(x)dx \quad \big(g \text{ هي غير سالبة}\big) \\ &\geq r\int_{\{x:g(x)\geq r\}}^{\infty} f_X(x)dx \\ &= rP(g(X) \geq r) \qquad \big(\text{تعريف}\big) \end{aligned}$$

ومنه

$$P(g(X) \geq r) \leq \frac{E\big(g(X)\big)}{r}$$

مثال 1.6. (**توضيح مساواة تشيبيتشيف**):

أكثر استعمالات مساواة تشيبيتشيف يكون فيها المتوسط والتباين. لتكن $g(X) = (x-u)^2/\sigma^2$، أين $u = E(X)$ و $\sigma^2 = V(X)$. و من أجل كتابة ملائمة $r = t^2$.:

$$P\left(\frac{(x-u)^2}{\sigma^2} \geq t^2\right) \leq \frac{1}{t^2} E\left(\frac{(x-u)^2}{\sigma^2}\right) = \frac{1}{t^2}$$

ببعض التبسيطات الجبرية نتحصل على:

$$P(|X-\mu| \geq t\sigma) \leq \frac{1}{t^2}$$

ونظيرتها:

$$P(|X-\mu| < t\sigma) \geq 1 - \frac{1}{t^2} \dots \dots (6.2)$$

التي تعطي مجال أو حد عام لانحراف $|X-\mu|$ عن σ. مثلا، بأخذ $t=2$، نتحصل على:

$$P(|X-\mu| \geq 2\sigma) \leq \frac{1}{2^2} = 0{,}25,$$

$$P(|X-\mu| < 2\sigma) \geq 1 - \frac{1}{t^2} = 0{,}75 = 75\%$$

لدينا على الأقل 75% من الحظ (الاحتمال) أن يقع متغير عشوائي بعيد عن متوسطه على الأكثر ب 2σ (بغض النظر عن التوزيع الاحتمالي ل X).

مثال 2.6. (**مساواة الاحتمال الطبيعي**):

ليكن Z متغير طبيعي معياري، إذا:

$$P(|Z| \geq t) \leq \sqrt{\frac{2}{\pi}} \frac{e^{-t^2/2}}{t}, \quad 0 < t \text{ أجل كل } \dots \dots (6.3)$$

بمقارنتها بمساواة تشيبيتشيف. من أجل $t=2$، تعطي $P(|Z| \geq t) \leq 0{,}25$ لكن $\sqrt{(2/\pi)}e^{-2}/2 = 0{,}054$، تحسن كبير.

لإثبات (3.6)، نكتب:

$$P(Z \geq t) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-x^2/2} dx$$

$$\leq \frac{1}{\sqrt{2\pi}} \int_t^\infty \frac{x}{t} e^{-x^2/2} dx \quad \text{كون } \frac{x}{t} > 1 \text{ من أجل } x > t$$

$$= \frac{1}{\sqrt{2\pi}} \frac{e^{-t^2/2}}{t}$$

وباستعمال العلاقة $P(|Z| \geq t) = 2P(Z \geq t)$. حد أدنى ل $P(|Z| \geq t)$ يمكن الحصول عليه بطريقة مماثلة.

توجد مساواة احتمالية عديدة أخرى والتي أغلبها مماثلة في المبدأ لمساواة تشيبيتشيف مثل العلاقة:

$$P(X \geq a) \leq e^{-at}M_X(t),$$

لكن هذه المساواة تتطلب وجود دالة توليد العزوم. علاقات مساواة أخرى أكثر دقة من مساواة تشيبيتشيف موجودة لكن تتطلب فرضيات أكثر.

6.ب. علاقات الثوابت:

في هذه الفقرة نقدم عينة من مختلف علاقات الثوابت التي يمكن ان تكون مفيدة ليس فقط في تأسيس بعض النظريات لكن أيضا للقيام بالحسابات المتعددة. مجموعة كاملة من علاقات الثوابت يمكن ان يرى إليها على أنها "علاقات تراجعية"، حيث رأينا القليل منها سابقا. نذكر أنه إذا كان $X \sim P(\lambda)$، فإن:

$$P(X = x+1) = \frac{\lambda}{x+1}P(X = x) \ldots\ldots\ldots (6.4)$$

تسمح لنا هذه العلاقة بحساب احتمالات توزيع بواسون تراجعيا بداية من $P(X = 0) = e^{-\lambda}$. أيضا نجد العلاقات من شكل العلاقة (2.6) من أجل كل التوزيعات المتقطعة. أحيانا توجد بعض الصيغ المختلفة البسيطة بالنسبة للتوزيعات المستمرة.

نظرية 6.1.:

ليكن $X_{\alpha,\beta}$ متغير عشوائي يتبع توزيع غاما $G(\alpha,\beta)$ ذو دالة كثافة $f(x/\alpha,\beta)$، أين $\alpha > 1$. من أجل أي ثابتين موجبين a و b يكون لدينا:

$$P\left(a < X_{\alpha,\beta} < b\right) = \beta[f(a/\alpha,\beta) - f(b/\alpha,\beta)] + P\left(a < X_{\alpha-1,\beta} < b\right) \ldots\ldots. (6.5)$$

البرهان:

لدينا من تعريف توزيع غاما:

$$P\left(a < X_{\alpha,\beta} < b\right) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}}\int_a^b x^{\alpha-1}e^{-x/\beta}\,dx$$

$$= \frac{1}{\Gamma(\alpha)\beta^{\alpha}}\left[\left[-x^{\alpha-1}\beta e^{-x/\beta}\right]_a^b + \int_a^b (\alpha-1)x^{\alpha-2}\beta e^{-x/\beta}\,dx\right]$$

حيث قمنا بالتكامل بالتجزئة وذلك بوضع $u = x^{\alpha-1}$ و $dv = e^{-x/\beta}dx$.

نجد:

$$P\left(a < X_{\alpha,\beta} < b\right) = \beta[f(a/\alpha,\beta) - f(b/\alpha,\beta)] + \frac{(\alpha-1)}{\Gamma(\alpha)\beta^{\alpha-1}}\int_a^b x^{\alpha-2}e^{-x/\beta}\,dx$$

باستعمال حقيقة أن: $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$ يصبح:

$$P\left(a < X_{\alpha,\beta} < b\right) = \beta[f(a/\alpha,\beta) - f(b/\alpha,\beta)] + \frac{1}{\Gamma(\alpha-1)\beta^{\alpha-1}}\int_a^b x^{\alpha-2}e^{-x/\beta}\,dx$$

نلاحظ أن الجزء الأخير من الطرف الأيمن من هده العلاقة ما هو إلا $P\left(a < X_{\alpha-1,\beta} < b\right)$. وبالتالي نجد في النهاية:

$$P\left(a < X_{\alpha,\beta} < b\right) = \beta[f(a/\alpha,\beta) - f(b/\alpha,\beta)] + P\left(a < X_{\alpha-1,\beta} < b\right)$$

وهو المطلوب.

نتيجة 1.6. (نتيجة ***Stein***):

ليكن $X \sim N(\mu,\sigma^2)$ ولتكن g دالة قابلة للاشتقاق وتحقق $E|g'(X)| < \infty$. إذا:

$$E[g(X)(X-\theta)] = \sigma^2 E[g'(X)]$$

البرهان:

$$E[g(X)(X-\theta)] == \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} g(x)(x-\theta)e^{-(x-\theta)^2/(2\sigma^2)}\,dx$$

بالتكامل بالتجزئة مع وضع $u = g(x)$ و $dv = (X-\theta)e^{-(x-\theta)^2/(2\sigma^2)}dx$ نتحصل على:

$$\begin{aligned} E[g(X)(X-\theta)] &= \\ &= \frac{1}{\sqrt{2\pi}\sigma}\left[\left[-\sigma^2(x)(x-\theta)e^{-(x-\theta)^2/(2\sigma^2)}\right]_{-\infty}^{+\infty} \right. \\ &\quad \left. + \sigma^2\int_{-\infty}^{\infty} g'(x)e^{-(x-\theta)^2/(2\sigma^2)}\,dx\right] \end{aligned}$$

الشرط المحقق من g كاف ليجعل العنصر الأول من الطرف الأيمن معدوم (يساوي 0) ويبقى العنصر الثاني الذي يكتب $\sigma^2 E[g'(X)]$.

مثال 3.6 (**عزوم التوزيع الطبيعي ذات الدرجة العالية**):

نتيجة ستين ($Stein$) تجعل من حساب العزوم ذات الدرجة العالية سهلا. مثلا، ليكن $X \sim N(\mu,\sigma^2)$، إذا:

$$\begin{aligned} E(X^3) &= E\left((X^2)(X-\theta+\theta)\right) \\ &= E\left((X^2)(X-\theta)\right) + \theta E(X^2) \\ &= 2\sigma^2 E(X) + \theta E(X^2) \qquad (g(x) = x^2, g'(x) = 2x\,) \\ &= 3\theta\sigma^2 + \theta^3\,. \end{aligned}$$

تكاملات بالتجزئة مماثلة لعلاقات الثوابت توجد من أجل العديد من التوزيعات (*Hudson 1978*). يمكننا الحصول أيضا على علاقات ثوابت مفيدة من خلال استغلال خصائص بعض التوزيعات الخاصة، مثلما توضحه النظرية التالية.

نظرية 2.6.:

ليكن χ_p^2 يشير إلى متغير عشوائي يخضع لتوزيع كاي تربيع بدرجة حرية p. من أجل أي دالة $h(x)$،

$$E\left(h\left(\chi_p^2\right)\right) = pE\left(\frac{h\left(\chi_{p+2}^2\right)}{\chi_{p+2}^2}\right) \dots\dots\dots (6.6)$$

شرط أن يكون الأمل موجودا (منتهي).

البرهان:

عبارة " شرط أن يكون الأمل موجودا " هي طريقة لتفادي وضع شروط على الدالة h. عموما، دوال عادية تحقق العلاقة (4.6). لدينا من العلاقة (6.6):

$$E\left(h\left(\chi_p^2\right)\right) = \frac{1}{\Gamma(p/2)2^{(p/2)}}\int_0^{\infty} h(x)x^{(p/2)-1}e^{-x/2}dx$$

$$= \frac{1}{\Gamma(p/2)2^{(p/2)}}\int_0^{\infty}\left(\frac{h(x)}{x}\right)x^{((p+2)/2)-1}e^{-x/2}dx$$

أين قمنا الطرف تحت التكامل ب (x/x). الآن نكتب:

$$\Gamma(p/2)2^{(p/2)} = \frac{\Gamma((p+2)/2)2^{(p+2)/2}}{p}$$

ويكون لدينا:

$$E\left(h\left(\chi_p^2\right)\right) = \frac{p}{\Gamma((p+2)/2)2^{(p+2)/2}}\int_0^{\infty}\left(\frac{h(x)}{x}\right)x^{((p+2)/2)-1}e^{-x/2}dx$$

$$= pE\left(\frac{h\left(\chi_{p+2}^2\right)}{\chi_{p+2}^2}\right).$$ وهو المطلوب

بعض حسابات العزوم تصبح سهلة مع العلاقة (6.6). مثلا، متوسط χ_p^2 هو:

$$E\left(\chi_p^2\right) = pE\left(\frac{\chi_{p+2}^2}{\chi_{p+2}^2}\right) = pE(1) = p,$$

والعزم من الدرجة الثانية هو:

$$E\left(\left(\chi_p^2\right)^2\right) = pE\left(\frac{\left(\chi_{p+2}^2\right)^2}{\chi_{p+2}^2}\right) = pE\left(\chi_{p+2}^2\right) = p(p+2),$$

إذا:

$$V\left(\chi_{p+2}^2\right) = p(p+2) - p^2 = 2p.$$

نختم هذا العنصر من علاقات الثوابت ببعض العلاقات المتقطعة المماثلة للعلاقة الأخيرة. الصيغة a العامة للعلاقتين في النظرية (8.6) التالية ترجع إلى هوانغ *Hwang* (1982).

نظرية 3.6. (نظرية هوانغ ***Hwang***):

لتكن $g(x)$ دالة مع $-\infty < E\big(g(X)\big) < \infty$ و $-\infty < g(-1) < \infty$. إذا:

أ. إذا كان $X \sim P(\lambda)$،

$$E\big(\lambda g(X)\big) = E\big(Xg(X-1)\big) \dots\dots\dots (6.7).$$

ب. إذا كان $X \sim B(r,p)$،

$$E\big((1-p)g(X)\big) = E\left(\frac{X}{r+X-1}g(X-1)\right) \dots\dots\dots (6.8).$$

البرهان:

نبرهن العلاقة "أ" ونترك العلاقة "ب" كتمرين. لدينا:

$$\begin{aligned} E\big(\lambda g(X)\big) &= \sum_{x=0}^{\infty} \lambda g(x) \frac{e^{-\lambda}\lambda^{x}}{x!} \\ &= \sum_{x=0}^{\infty} g(x) \frac{e^{-\lambda}\lambda^{x+1}}{x!}\frac{(x+1)}{(x+1)} \\ &= \sum_{x=0}^{\infty} (x+1)g(x) \frac{e^{-\lambda}\lambda^{x+1}}{(x+1)!} \end{aligned}$$

نقوم الآن بتحويل مؤشر المجموع بكتابة "$y = x + 1$". كون x يتغير من 0 إلى ∞ فإن y يتغير من 1 إلى ∞. ويصبح:

$$\begin{aligned} E\big(\lambda g(X)\big) &= \sum_{y=1}^{\infty} yg(y-1) \frac{e^{-\lambda}\lambda^{y}}{y!} \\ &= \sum_{y=0}^{\infty} yg(y-1) \frac{e^{-\lambda}\lambda^{y}}{y!} \\ &= E\big(Xg(X-1)\big) \end{aligned}$$

كون هذا المجموع الأخير هو أمل رياضي ل $P(\lambda)$.

استعمل "هوانغ" هذه العلاقة بطريقة مماثلة لستين "Stein"، لبرهنة نتائج متعلقة بتقديرات متعددة كما أن لعلاقة هوانغ تطبيقات أخرى، خاصة لحساب العزوم.

مثال 4.6 (**عزوم هيغر لتوزيع بواسون**):

من أجل $X \sim P(\lambda)$، نضع $g(x) = x^2$ وباستعمال العلاقة (5.6) نجد:

$$E(\lambda X^2) = E(X(X-1)^2) = E(X^3 - 2X^2 + X).$$

لكن العزم الثالث لتوزيع $P(\lambda)$ هو:

$$\begin{aligned} E(X^3) &= \lambda E(X^2) + 2E(X^2) - E(X) \\ &= \lambda(\lambda + \lambda^2) + 2(\lambda + \lambda^2) - \lambda \\ &= \lambda^3 + 3\lambda^2 + \lambda \end{aligned}$$

من توزيع ثنائي الحد السالب، المتوسط يمكن حسابه بأخذ $g(x) = r + x$ في العلاقة (8.6):

$$E\big((1-p)(r+x)\big) = E\left(\frac{X}{r+X-1}(r+x-1)\right) = E(X).$$

ببعض الترتيبات نتحصل على:

$$\big(E(X)\big)\big((1-p)-1\big) = -r(1-p)$$

ومنه:

$$E(X) = \frac{r(1-p)}{p}.$$

عزوم أخرى يمكن حسابها بنفس الطريقة.

7. **تمارين:**

1.7. أوجد صيغ ل $E(X)$ و $V(X)$ إذا كان المتغير العشوائي X يتبع التوزيع أحادي الشكل المتقطع $U(N_0;N_1)$ بحيث N_0 و N_1 عددين طبيعيين و $(N_0 < N_1)$.

2.7. يستقبل مصنع من ممون طلبية من 100 وحدة حيث تعتبر الطلبية مرفوضة إذا كانت تحتوي على أكثر من 5 وحدات فاسدة. يقوم المصنع بالاختيار العشوائي ل K وحدة من الطلبية من أجل المعاينة بحيث تقبل الطلبية إذا لم تحتوي العينة المختارة ولا وحدة فاسدة:

أ- ما هو حجم العينة K الكافي لضمان أن يكون احتمال أن يقبل المصنع طلبية غير مقبولة أقل من *0,1*؟

ب- نفرض أن المصنع قرر أن يقبل الطلبية إذا كان على الأكثر وحدو واحدة فاسدة في العينة المختارة. ما هو حجم العينة K الضروري لضمان أن يكون احتمال أن يقبل المصنع طلبية غير مقبولة أقل من *0,1*؟

3.7. تدفق حركة المرور في إحدى الطرق يمكن أن يمثل أو ينمذج بتتابع أحداث برنولي وذلك بافتراض أن احتمال مرور سيارة في أي ثانية هو ثابت p ولا يوجد تداخل بين مرور السيارات في الثواني المختلفة. لو نعامل الثواني على أساس وحدات زمنية غير قابلة للتجزئة (حوادث) حيث يمكننا تطبيق نموذج برنولي. نفرض أن راجل يمكنه قطع الطريق فقط في حالة عدم مرور سيارة في الثواني 3 القادمة. أوجد احتمال أن الراجل له أن ينتظر بالضبط 4 ثواني قبل أن يبدأ بقطع الطريق.

4.7 دواء معروف بفعاليته في *80%* من حالات استعماله. دواء جديد معادل تم اختباره على 100 مريض ووجد أنه فعال في 85 حالة. هل الدواء الجديد أفضل؟ (مساعدة: احسب احتمال مشاهدة 85 أو أكثر حالة فعالية بافتراض أن الدواءين، القديم والجديد، هما فعليا متكافئين).

5.7 عدد كبير من الحشرات متوقع أن ينجذب لنوع معين من الأزهار النباتية. تاجر مبيدات يدعي أنها فعالة بنسبة *99%*. نفرض أن 2000 حشرة غزت حديقة أزهار بعد استعمال المبيد و ليكن X: عدد الحشرات الناجية.

أ- ما هو التوزيع الاحتمالي الذي قد يمثل نموذج لهذه التجربة؟

ب- اكتب، دون حساب ذلك، عبارة احتمال لن يكون أقل من 100 حشرة ناجية، مستعملا النموذج المقترح في 1.

ت- أعط قيمة تقريبية للاحتمال المطلوب في 2.

6.7 ليكن عدد حبات الشوكولاتة في نوع معين من الحلويات يتبع توزيع بواسون. نرغب في أن يكون احتمال أن تكون قطعة حلوى مختارة عشوائيا بها على الأقل قطعتي شوكولاتة أكبر من *0,99*. أوجد أصغر قيمة لمتوسط هذا التوزيع لتحقق هذا الاحتمال.

7.7 قاعتي دور سينما يتنافسان على 1000 زبون. بافتراض أن كل زبون يختار بينهما باستقلالية وبدون اعتبار. ليكن N عدد المقاعد في كل قاعة.

أ- باستعمال نموذج أو توزيع ثنائي الحد، أوجد عبارة ل N التي تضمن أن يكون احتمال رفض زبون بسبب امتلاء القاعة أقل من *1%*.

ب- استعمل التقريب الطبيعي لإعطاء قيمة عددية ل N.

8.7 يمكن تقريب توزيع فوق الهندسي إما بتوزيع ثنائي الحد أو بتوزيع بواسون. (بالطبع، يمكن تقريبه بتوزيعات أخرى أيضا). ليكن X متغير عشوائي الذي يخضع للتوزيع الفوق الهندسي بحيث:

$$P(X = x/N\,,M,K) = \frac{\binom{x}{M}\binom{k-x}{N-M}}{\binom{k}{N}}, x = 0,1,2,\ldots,K$$

أ- بين أنه لما $N \to \infty$، $M \to \infty$ و $M/N \to p$ فإنه

$$P(X = x/N, M, K) = \binom{x}{K} p^x (1-p)^{k-x}; \quad x = 0,1,2, \dots, K$$

ب– استعمل خاصية أن توزيع ثنائي الحد يمكن تقريبه بتوزيع بواسون لإظهار أنه عندما $N \to \infty$ ،$M \to \infty$ ، $K \to \infty$ ، $M/N \to 0$ و$KM/N \to \lambda$ فإنه يكون:

$$P(X = x/N, M, K) = \frac{e^{-\lambda}\lambda^x}{x!}; \; x = 0,1,2, \dots$$

ت– تحقق من التقريب في 2 مباشرة دون استعمال تقريب بواسون إلى توزيع ثنائي الحد.

9.7 نفرض أن $X \sim B(n;p)$ وليكن $Y \sim NB(r;p)$. بين أن $F_X(r-1) = 1 - F_Y(n-r)$.

10.7 التوزيع المتقطع الناقص هو التوزيع الذي من أجله فئة لا يمكن مشاهدتها وتقصى من فضاء أو مجموعة القيم. بالخصوص، إذا كان X يأخذ الرتب $0,1,2, \dots$ و الفئة 0 لا يمكن مشاهدتها (الحالة الغالبة)، المتغير العشوائي الناقص X_T له دالة كثافة احتمالية:

$$P(X_T = x) = \frac{P(X = x)}{P(X > 0)}; x = 1,2, \dots$$

أوجد دالة الكثافة الاحتمالية، المتوسط والتباين للمتغير العشوائي الناقص X_T في الحالتين:

أ. $X \sim P(\lambda)$

ب. $X \sim BN(r;p)$

11.7 لاحظنا أن توزيع بواسون $P(\lambda)$ هو نهاية لتوزيع ثنائي الحد السالب $BN(r;p)$ لما $r \to \infty$، $p \to 1$ و $r(1-p) \to \lambda$، بين أنه تحت هذه الشروط، دالة الكثافة الاحتمالية لتوزيع ثنائي الحد السالب تقارب تلك لتوزيع بواسون.

12.7 تحقق من العلاقتين أسفله وذلك بالرجوع إلى دالة غاما المعطاة فيما قبل:

ت. $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$،

ث. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

13.7 ضع علاقة مماثلة ل (18.3.) لتوزيع غاما. إذا كان $X \sim G(\alpha;\beta)$، إذا من أجل كل عدد موجب ثابت v:

$$E(X^v) = \frac{\beta^v \Gamma(v+\alpha)}{\Gamma(\alpha)}$$

14.7 توجد علاقة مهمة بين توزيعي ثنائي الحد السالب و غاما. ليكن Y متغير عشوائي ثنائي الحد ذو معلمة r و p، حيث تمثل p احتمال النجاح أو التحقق. بين أنه لما $p \to 0$ فإن دالة الكثافة الاحتمالية للمتغير العشوائي (pY) تتقارب نحو توزيع غاما ذو المعلمتين r و 1.

15.7 بين أن:

$$\int_z^{\infty} \frac{1}{\Gamma(\alpha)} z^{\alpha-1} e^{-z} dz = \sum_{y=0}^{\alpha-1} \frac{x^y e^{-x}}{y!}; \; \alpha = 1,2,3, \dots$$

(ملاحظة: نستعمل التكامل بالتجزئة.) عبر عن هذه الصيغة كعلاقة احتمالية بين متغيرين عشوائيين لبواسون وغاما.

16.7 ليكن المتغير العشوائي X ذو دالة الكثافة الاحتمالية:

$$f(x) = \frac{2}{\sqrt{2\pi}} e^{-x^2/2};\ \ 0 < x < \infty.$$

ج. أوجد متوسط وتباين X. (يسمى أحيانا هذا **التوزيع الطبيعي المطوي**).

ح. إذاكان X يتبع التوزيع الطبيعي المطوي، أوجد التحويل $Y = g(x)$ وقيم α و β بحيث يكون $Y \sim G(\alpha;\beta)$.

17.7 اكتب التكامل الذي يعرف دالة العزوم الاحتمالية لدالة الكثافة الاحتمالية:

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$$

هل التكامل منتهي؟ (هل كنت تتوقع أن يكون كذلك؟).

18.7 من أجل التوزيعات التالية، تحقق من صيغ $E(X)$ و $V(X)$:

أ- تحقق من $V(X)$ لما $X \curvearrowright P(\lambda)$ (ملاحظة: احسب $E\big(X(X-1)\big) = E(X^2) - E(X)$).

ب- تحقق من $V(X)$ لما $X \curvearrowright NB(r;p)$،

ت- تحقق من $V(X)$ لما $X \curvearrowright G(\alpha;\beta)$،

ث- تحقق من $E(X)$ لما $X \curvearrowright B(\alpha;\beta)$،

ج- تحقق من $E(X)$ و $V(X)$ لما $X \curvearrowright DE(\mu;\sigma)$.

19.7 توزيع باريتو Pareto ذو المعالم α و β، لديه دالة الكثافة الاحتمالية:

$$f(x) = \frac{\beta\alpha^\beta}{x^{\beta+1}};\ \ \alpha < x < \infty;\ \alpha > 0;\ \beta > 0.$$

أ- تحقق من أن $f(x)$ هي دالة كثافة احتمالية،

ب- احسب متوسط وتباين هذا التوزيع،

ت- بين أن تباين هذا التوزيع غير موجود لما $\beta \le 2$.

20.7 أغلب التوزيعات "المسماة" تمثل حالة خاصة من التوزيعات العامة المشار إليها. من أجل التوزيعات المسماة الآتية، استنتج صيغة دالة الكثافة الاحتمالية، تحقق من أنها كذلك، واحسب المتوسط والتباين:

أ- إذاكان $X \curvearrowright E(\beta)$ فإن $Y = X^{1/\gamma} \curvearrowright W(\gamma;\beta)$ أين $\gamma > 0$ هو ثابت،

ب- إذاكان $X \curvearrowright E(\beta)$ فإن $Y = 2X/\beta \curvearrowright R$ بحيث R: تعني توزيع رايليخ *Rayleigh*،

ت- إذاكان $X \curvearrowright G(\mathrm{a};\mathrm{b})$ فإن $Y = 1/X \curvearrowright IG(\mathrm{a};\mathrm{b})$ بحيث IG: تعني توزيع غاما المعكوس *Inverted Gamma*،

ث- إذاكان $X \curvearrowright G\left(\frac{3}{2};\beta\right)$ فإن $Y = (X/\beta)^{1/2} \curvearrowright M$ بحيث M: تعني توزيع غاما ماكسويل *Maxwell*،

ج- إذاكان $X \curvearrowright E(1)$ فإن $Y = \alpha - \gamma lnX \curvearrowright Gumbel(\alpha;\gamma)$ أين $-\infty < \alpha < \infty$ و $\gamma > 0$, (يعرف توزيع غامبل أيضا بتوزيع القيمة العظمى).

21.7 ليكن المتغير العشوائي T الذي يمثل مدة حياة لعنصر ما (قد يكون مركب إلكتروني أو أي شيء آخر). الدالة العشوائية $h_T(t)$ التابعة للمتغير العشوائي T هي معرفة ب:

$$h_T(t) = \lim_{\delta \to 0} \frac{P(t \leq T < t + \delta \,/\, T \geq t)}{\delta}.$$

هنا يمكننا أن نفسر $h_T(t)$ على أنها نسبة التغير في احتمال أن العنصر يعيش فترة قصيرة بعد المدة t علما أنه عاش حتى المدة t. بين أنه إذا كان T متغير مستمر فإن:

$$h_T(t) = \frac{f_T(t)}{1 - F_T(t)} = -\frac{d}{dt} ln\big(1 - F_T(t)\big).$$

22.7 تحقق من أن دوال الكثافة الاحتمالية التالية هي الدوال العشوائية المشار إليها:

أ- إذا كان $T \sim E(\beta)$ فإن $h_T(t) = 1/\beta$،

ب- إذا كان $T \sim W(\gamma; \beta)$ فإن $h_T(t) = (\gamma/\beta)t^{\gamma-1}$،

ت- إذا كان $T \sim Logistic(\mu; \beta)$ فإن:

$$F_T(t) = \frac{1}{1 + e^{-(t-\mu)/\beta}}$$

إذا $h_T(t) = (1/\beta)F_T(t)$

23.7 من أجل العائلات التالية، بين التي دوال كثافتها الاحتمالية هي وحيدة المنوال:

أ- $U(a; b)$،

ب- $G(\alpha; \beta)$،

ت- $N(\mu; \sigma^2)$،

ث- $B(\alpha; \beta)$,

24.7 بين أن كل من العائلات التالية هي عائلة أسية:

أ- عائلة طبيعية لها معالم μ و σ معلومة،

ب- عائلة غاما لها معالم α و β معلومة أو كلاهما مجهولة،

ت- عائلة بيتا لها معالم α و β معلومة أو كلاهما مجهولة،

ث- عائلة بواسون،

ج- عائلة ثنائي الحد السالب ذا ت معلمة r معلومة، $0 < p < 1$.

25.7 من أجل كل عائلة في التمرين 24.7، اشرح فضاء المعالم الطبيعية.

26.7 استعمل علاقات ثوابت النظرية (1.4.) ل:

أ- حساب تباين متغير عشوائي ثنائي الحد،

ب- حساب متوسط وتباين متغير عشوائي بيتا $B(a; b)$.

27.7 لافي هذا التمرين سوف نبرهن النظرية (1.4.):

أ- ابدأ من المساواة

$$\int f(x/\theta)=h(x)c(\theta)exp\left(\sum_{i=1}^{k}w_i(\theta)t_i(x)\right)dx=1,$$

اشتق الطرفين ورتب الأطراف للحصول على العلاقة (3.4.). (العلاقة $\frac{d}{dx}lng(x)=g'(x)/g(x)$ مساعدة لإيجاد لذلك).

ب- اشتق العلاقة أعلاه مرة ثانية ثم رتب العناصر لتحصل على العلاقة (4.4.). (العلاقة $\frac{d^2}{dx^2}lng(x)=(g''(x)/g(x))-(g'(x)/g(x))^2$ مساعدة لإيجاد لذلك).

28.7

أ- إذا كان بالإمكان كتابة العائلة الأسية من الصيغة (7.4.)، بين أن علاقات ثوابت النظرية (1.4.) تبسط إلى:

$$E\left(t_j(x)\right)=-\frac{\partial}{\partial\eta_j}lngc^*(\eta),$$

$$V\left(t_j(x)\right)=-\frac{\partial^2}{\partial\eta_j^2}lngc^*(\eta).$$

ب- استعمل العلاقة هذه لحساب المتوسط والتباين لمتغير عشوائي لتوزيع غاما $G(a;b)$.

29.7 من أجل كل العائلات التالية:

- تأكد أنها عائلة أسية،
- مثل منحنى فضاء المعالم المنحني.

أ- $N(\theta;\theta)$

ب- $N(\theta;a\theta^2)$، a معلوم،

ت- $G(\alpha;1/\alpha)$،

ث- $f(x/\theta)=Ce^{(-(x-\theta)^4)}$، C هو ثابت معياري.

30.7 لاحظنا في المثال (5.4.) أن التقريبات الطبيعية يمكن أن تنتج عن العائلات الأسية المنحنية. من أجل التقريبات الطبيعية التالية:

أ- تقريب بواسون: $\bar{X}\sim N(\lambda;\lambda/n)$،

ب- تقريب ثنائي الحد: $\bar{X}\sim N(p;p(1-p)/n)$ ،

ت- تقريب ثنائي الحد السالب: $\bar{X}\sim N(r(1-p)/n;r(1-p)/np^2)$.

ث- مثل منحنى فضاء المعالم المنحني.

31.7

أ- العائلة الطبيعية التي تقرب توزيع بواسون يمكن أيضا أن تمثل ب $N\left(e^\theta;e^\theta\right)$ ، أين $-\infty<\theta<\infty$. مثل بيانيا فضاء المعالم، ثم قارن ذلك بالتقريب في التمرين 30.7-أ-.

ب- بافتراض أن $X\sim G(\alpha;\beta)$ وباعتبار $E(X)=\mu$، مثل بيانيا فضاء المعالم،

ت- بافتراض أن $X_i \sim G(\alpha_i;\beta_i); i=\overline{1,n}$ و $E(X_i)=\mu$، صف فضاء المعالم $(\alpha_1,\dots,\alpha_n,\beta_1,\dots,\beta_n)$.

32.7 لتكن دالة الكثافة الاحتمالية $f(X) = \frac{63}{4}(x^6 - x^8);\ -1 < x < 1$. مثل بيانيا، وفي نفس المعلم، $(1/\sigma)f((x-\mu)/\sigma)$ من أجل كل حالة من الحالات التالية:

أ- $\mu = 0$ و $\sigma = 1$،

ب- $\mu = 3$ و $\sigma = 1$،

ت- $\mu = 3$ و $\sigma = 2$،

33.7 بين أنه إذا كانت $f(X)$ دالة كثافة احتمالية، متناظرة حول 0، فإنه يكون μ الوسيط لدالة الكثافة الاحتمالية $(1/\sigma)f((x-\mu)/\sigma);\ -\infty < x < \infty$.

34.7 ليكن Z متغير عشوائي و $f(z)$ دالة كثافته الاحتمالية. حدد أو عرف العدد z_α حتى يحقق العلاقة:

$$\alpha = P(Z > z_\alpha) = \int_{z_\alpha}^{\infty} f(z)\, dz.$$

بين أنه إذا كان X متغير عشوائي دالة كثافته الاحتمالية $(1/\sigma)f((x-\mu)/\sigma)$ و $x_\alpha = \sigma z_\alpha + \mu$ فإن $\alpha = P(X > x_\alpha)$.

35.7 لنكن عائلة كوشي المعرفة في الفقرة (3.ج) يمكن توسعة هذه العائلة إلى عائلة الموضع–القياس التي تكون دالة كثافتها الاحتمالية من الشكل:

$$f(x/\mu,\sigma) = \frac{1}{\sigma\pi\left(1 + \left(\frac{x-\mu}{\sigma}\right)^2\right)};\ -\infty < x < \infty.$$

نعلم أن المتوسط والتباين غير موجودين لتوزيع كوشي. إذا المعلمتين μ و σ^2 لا تمثلان المتوسط والتباين لكن لهما معنى مهم بالنسبة للتوزيع.

بين أنه إذا كان X متغير عشوائي يخضع لتوزيع كوشي ذو معلمتين μ و σ فإن:

أ- μ هو الوسيط الحسابي لتوزيع X أي $P(X \geq \mu) = P(X \leq \mu) = \frac{1}{2}$،

ب- $\mu + \sigma$ و $\mu - \sigma$ هي الربعيات لتوزيع X بحيث $P(X \geq \mu + \sigma) = P(X \leq \mu - \sigma) = \frac{1}{4}$

(مساعدة: برهن أولا من أجل $\mu = 0$ و $\sigma = 1$ ثم استعن بالتمرين 7.ذ.ذ) .

36.7 لتكن $f(x)$ أي دالة كثافة احتمالية بمتوسط μ و تباين σ^2. بين كيفية تشكيل عائلة موضع–قياس تعتمد على $f(x)$ بحيث دالة الكثافة الاحتمالية المعيارية للعائلة، نرمز لها $f^*(x)$، لها متوسط 0 وتباين 1.

37.7 عائلة دوال كثافة تراكمية $\{F(x/\theta); \theta \in \Theta\}$ هي متزايدة عشوائيا بدلالة θ إذا كان $\theta_1 > \theta_2 \Leftarrow F(x/\theta_1)$ عشوائيا أكبر من $F(x/\theta_2)$.

38.7 ارجع إلى التمرين 37.7 من أجل تعريف عائلة متزايدة عشوائيا.

أ- بين أن عائلة موضع هي متزايدة عشوائيا بدلالة معلمة موضعها،

ب- بين أن عائلة قياس هي متزايدة عشوائيا بدلالة معلمة قياسها إذا كان فضاء العينات هو $[0;\infty[$.

39.7 عائلة دوال كثافة تراكمية $\{F(x/\theta); \theta \in \Theta\}$ هي متناقصة عشوائيا بدلالة θ إذا كان $\theta_1 > \theta_2 \Leftarrow F(x/\theta_2)$ عشوائيا أكبر من $F(x/\theta_1)$. (انظر التمرين 37.7 و 38.7).

أ- بين أنه إذا كان $X \backsim F_X(x/\theta)$، أين فضاء العينات ل X هو $]0;\infty[$ و $F_X(x/\theta)$ هي متزايدة عشوائيا بدلالة θ، إذا $F_Y(y/\theta)$ هي متناقصة عشوائيا بدلالة θ، أين $Y = 1/X$.

ب- بين أنه إذا كان $X \backsim F_X(x/\theta)$، أين $F_X(x/\theta)$ هي متزايدة عشوائيا بدلالة θ و $\theta > 0$، إذا $F_X\left(x/\frac{1}{\theta}\right)$ هي متناقصة عشوائيا بدلالة θ.

40.7 من أجل كل متغير عشوائي X أين يكون $E(X^2)$ و $E(|X|)$ موجودين، بين أن قيمة $P(|X| \geq b)$ لا تتجاوز لا $E(X^2/b^2)$ ولا $E(|X|/b)$، أين b هو عدد ثابت موجب. إذا كان $f(x) = e^{-x}$ من أجل $x > 0$ ، بين أن حدا يكون أفضل لما $b = 3$ والآخر لما $b = \sqrt{2}$.

41.7 ليكن المتغير العشوائي X ذو دالة توليد العزوم $M_X(t)$، $-h < t < h$.

أ- برهن أن $P(X \geq a) \leq e^{-at}M_X(t)$، $0 < t < h$. (برهان مماثل لذلك المستعمل لبرهنة علاقة تشيبي شيف.)،

ب- برهن كذلك أن، $P(X \leq a) \leq e^{-at}M_X(t)$ $-h < t < 0$،

ت- حالة خاصة من 1 هي $P(X \geq 0) \leq E(e^{tx})$ من أجل كل $t \geq 0$ التي من أجلها تكون دالة توليد العزوم معرفة. ما هي الشروط العامة للدالة $h(t,x)$ حتى يكون $P(X \geq 0) \leq E\big(h(t,x)\big)$ من أجل كل $t \geq 0$ التي من أجلها $E\big(h(t,x)\big)$ معرفة؟ (في الجزء 1، $h(t,x) = e^{tx}$).

42.7 احسب $P(|X - \mu_X| \geq k\sigma_X)$ من أجل $X \backsim U(0;1)$ و $X \backsim E(\lambda)$ وقارن إجاباتك بحدود عدم مساواة تشيبي تشيف.

43.7 إذا كان Z متغير طبيعي معياري، برهن هذه العلاقة الموافقة لعدم المساواة في المثال (2.6.):

$$P(|Z| \geq t) \geq \sqrt{\frac{2}{\pi}}\frac{t}{1+t^2}e^{-t^2/2}.$$

44.7 برهن العلاقات التراجعية، مثل تلك المعطاة في المثال (1.6.)، من أجل التوزيعات، ثنائي الحد، ثنائي الحد السالب والفوق الهندسي.

45.7 برهن العلاقات التالية المماثلة لعلاقة ستين *Stein* ، مفترضا الشروط الملائمة على الدالة g.

أ- إذا كان $X \sim G(\alpha;\beta)$، إذا $E\big(g(X)(X - \alpha\beta)\big) = \beta E\big(Xg'(X)\big)$.

ب- إذا كان $X \sim B(\alpha;\beta)$، إذا $E\left(g(X)\left(\beta - (\alpha - 1)\frac{(1-X)}{X}\right)\right) = E\Big(\big((1-X)\big)g'(X)\Big)$.

برهن علاقة (ثابت) توزيع ثنائي الحد السالب المعطاة في النظرية (3.6.)، الجزء ب.

8. متنوعات:

8.أ. فرضيات بواسون:

نظرية 1.8:

من أجل $t \geq 0$، ليكن N_t متغير عشوائي طبيعي يتمتع بالخصائص التالية (مثلا N_t يمثل عدد القادمين في الفترة الزمنية من 0 إلى t):

أ- $N_0 = 0$، (عدم وجود قادمين في البداية)

ب- $s < t \Leftarrow N_s$ و $N_t - N_s$ هي مستقلة، (عدد القادمين في فترات مختلفة هو مستقل)

ت- N_s و $N_{t+s} - N_t$ هي متماثلة التوزيع، (عدد القادمين هو مرتبط فقط بطول الفترة)

ث- $\lim_{t\to 0} \frac{P(N_t=1)}{t} = \lambda$ (احتمال القدوم هي نسبية لطول الفترة، إذاكان هذا الأخير صغيرا)

ج- $\lim_{t\to 0} \frac{P(N_t>1)}{t} = \lambda$، (لا قادمين في آن واحد)

$$P(N_t = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!},$$

ونقول، $N_t \rightsquigarrow P(\lambda t)$.

يمكن أيضا تفسير الفرضيات على أنها تشرح أو تمثل سلوك أشياء فضائية (مثلا، حركة الحشرات)، الأمر الذي يسمح بتطبيق بواسون في التوزيعات الفضائية.

8.ب. حدود تشيبيتشيف *Chebychev and Beyond*:

غوش *Ghosh* و ميدن *Meeden* (1977) ناقشا حقيقة أن علاقة عدم المساواة لتشيبيتشيف هي أكثر عموما وغالبا لا يتم بلوغها. لو نرمز ب $\bar{X}_n$ ل متوسط المتغيرات العشوائية $X_1, X_2, \dots, X_n$، إذا حسب علاقة عدم المساواة لتشيبيتشيف يكون:

$$P(|\bar{X}_n - \mu| \geq k\sigma) \leq \frac{1}{nk^2}.$$

حيث برهنا النظرية التالية:

نظرية 2.8:

أ- إذاكان $n = 1$، فإن العلاقة هي محققة أو يتم بلوغها من أجل $k \geq 1$ و غير محققة أو لا يتم بلوغها من أجل $0 < k < 1$.

ب- إذاكان $n = 2$، فإن العلاقة هي محققة أو يتم بلوغها فقط من أجل $k = 1$.

ت- إذاكان $n \geq 3$، فإن العلاقة هي دائما غير محققة أو لايتم يتم بلوغها.

أمثلة معطاة من أجل حالات تحقق أو بلوغ علاقة عدم المساواة. أغلب مبرراتها التقنية تعتمد على علاقة عدم المساواة التالية، المعروفة ب علاقة عدم المساواة لماركوف *Markov's Inequality*:

إذاكان $P(Y \geq 0) = 1$ و $P(Y \leq 1) = 1$، إذا، من أجل أي $r > 0$،

$$P(Y \geq r) \leq \frac{E(Y)}{r}$$

مع مساواة إذا وفقط إذا كان $P(Y = r) = p = 1 - P(Y = 0)$، $0 < p < 1$.

علاقة عدم المساواة لماركوف يمكن أن تطبق على القيمة أو الكمية

$$Y = \frac{(\bar{X}_n - \mu)^2}{\sigma^2}$$

للحصول على النتائج أعلاه.

السبب الذي يجعل من علاقة عدم المساواة لتشيبيتشيف غير متماسكة هي عدم وضعها لقيود على التوزيع. مع قيود إضافية لأحادية المنوال، يمكننا الحصول على حدود أكثر قوة أو دقة ونتحصل على علاقات عدم المساواة لغوس و فيزوشونسكي–بيتيونين (*Gauss and Vysochanskii-Petunin*).

نظرية 4.8: (**عدم مساواة غوس** ***Gauss Inequality***)

ليكن $X \sim f$، أين f هي دالة وحيدة المنوال v، ونعرف $\tau^2 = E(X - v)^2$. إذا:

$$P(|X - v| > \varepsilon) \leq \begin{cases} \dfrac{4\tau^2}{9\varepsilon^2} & \text{من أجل } \varepsilon \geq \sqrt{4/3}\,\tau \\ 1 - \dfrac{\varepsilon}{\sqrt{3}\tau} & \text{من أجل } \varepsilon \leq \sqrt{4/3}\,\tau. \end{cases}$$

على الرغم من دقة (ضيق) الطرف مقارنة بطرف تشيبيتشيف، ارتباطها بالمنوال يحد من استعمالاتها. توسعة فيزوشونسكي–بيتيونين تزيح هذه المحدودية.

نظرية 5.8: (**عدم مساواة فيزوشونسكي–بيتيونين** ***Vysochanskii-Petunin Inequality***)

ليكن $X \sim f$، أين f هي وحيدة المنوال ونعرف $\xi^2 = E(X - \alpha)^2$ من أجل نقطة عادية α. إذا

$$P(|X - \alpha| > \varepsilon) \leq \begin{cases} \dfrac{4\xi^2}{9\varepsilon^2} & \text{من أجل } \varepsilon \geq \sqrt{8/3}\,\xi \\ \dfrac{4\xi^2}{9\varepsilon^2} - \dfrac{1}{3} & \text{من أجل } \varepsilon \leq \sqrt{8/3}\,\xi. \end{cases}$$

بيكلشايم *Pukelsheim* أشار أنه بأخذ أو وضع $\alpha = \mu = E(X)$ و $\varepsilon = 3\sigma$ أين $\sigma^2 = V(X)$ نتحصل على:

$$P(|X - \mu| > 3\sigma) \leq \frac{4}{81} < 0{,}05,$$

هذه العلاقة تسمى ب**قاعدة 3–الانحراف المعياري**، التي تعني أن الاحتمال يكون أقل من 5% أن يأخذ المتغير X قيمة أبعد من 3 أضعاف الانحراف المعياري عن متوسطه في المجتمع.

9. المتغيرات العشوائية الثنائية:

في العناصر السابقة، ناقشنا نماذج الاحتمالات وحساباتها التي تتعلق بمتغير عشوائي واحد. تسمى هذه **بنماذج أحادية المتغيرات**. في هذا العنصر، سنناقش نماذج الاحتمالات التي تشرك أو تتعلق بمتغيرين عشوائيين، **نماذج ثنائية المتغيرات**.

9.أ. التوزيعات المشتركة والتوزيعات الهامشية:

في الحالات التطبيقية، يكون من غير المعتاد القيام بتجربة يتم فيها مشاهدة قيم متغير عشوائي واحد فقط. مثلا، لتكن التجربة أو الدراسة التي تهدف إلى جمع معلومات حول الخصائص الصحية لمجتمع ما. ستكون تجربة متواضعة لو يتم فقط من خلالها الاهتمام أو جمع بيانات الوزن لشخص واحد. بالمقابل، أوزان عدة أشخاص في المجتمع يمكن قياسها وجمعها. هذه الأوزان المختلفة يمكن أن تمثل مشاهدات لمتغيرات عشوائية مختلفة، متغير لكل شخص يتم قياسه. مشاهدات متعددة يمكن أيضا أن تظهر بسبب تعدد الخصائص الفيزيائية الممكن قياسها ومشاهدتها لكل شخص من المجتمع. مثلا، الحرارة، الطول وضغط الدم زيادة على الوزن يمكن قياسها. هذه المشاهدات للخصائص المتعددة يمكن أيضا أن تنمذج كمشاهدات لمتغيرات عشوائية متعددة. لهذا، نحتاج إلى معرفة كيفية وصف واستعمال نماذج الاحتمالات التي تتعامل مع أكثر من متغير عشوائي في وقت واحد.

تعريف 1.9.:

متغير شعاعي ذو بعد n هو دالة من فضاء العينة S إلى فضاء إقليدي ذو بعد n، $\mathbb{R}^n$. نفرض، على سبيل المثال، أنه من أجل كل نقطة من فضاء العينة نربط زوج مرتب من الأعداد، $(x;y) \in \mathbb{R}^2$، أين $\mathbb{R}^2$ تشير إلى معلم. إذا قمنا بتعريف متغير عشوائي ثنائي البعد $(X;Y)$. المثال التالي يوضح ذلك.

مثال 1.9. (فضاء عينة لزهرة النرد):

لتكن التجربة العشوائية المتمثلة في رمي زهرتي نرد متزنتين. فضاء العينة لهذه التجربة يتشكل من 36 حالة أو نقطة متماثلة الاحتمال. مثلا، الحالة (3،3) تشير إلى حالة ظهور الوجه 3 في القطعتين، الحالة (4،1) تشير إلى حالة ظهور الوجه 4 في القطعة الأولى والوجه 1 في القطعة الثانية...إلخ. الآن، من أجل 36 حالة أو نقطة، نعرف المتغيرين X و Y بحيث:

X: مجموع القطعتين أو الزهرتين،

Y: القيمة المطلقة للفرق بين القطعتين.

من الحالة (3،3) نجد $X = 3 + 3 = 6$، و $Y = |3 - 3| = 6$ و من الحالة (4،1) نجد $X = 5$، و $Y = 4$، التي توافق أيضا قيم X و Y من الحالة (4،1). من أجل 36 حالة الممكنة يمكننا حساب قيم X و Y. بهذه الطريقة قمنا بتعريف شعاع متغيرين عشوائيين (X,Y).

بعد تعريفنا لشعاع المتغيرات العشوائية (X,Y)، يمكننا الآن مناقشة احتمالات الحوادث المعرفة بدلالة (X,Y). احتمالات الحوادث المعرفة بدلالة X و Y هي معرفة بدلالة الاحتمالات التي توافق الأحداث في فضاء العينة S. ما هو احتمال $P\left(X = 5 \text{ و } Y = 3\right)$؟ يمكننا التحقق أنه هناك فقط عنصرين أو نقطتين من العينة اللتين تحققان $X = 5$ و $Y = 3$، $(4,1)$ و $(1,4)$. أي أن الحادث

"$Y = 3$ و $X = 5$ " يتحقق إذا وفقط إذا تحقق الحادث $\{(4,1)\,,(1,4)\}$. كون كل عناصر 36 من العينة S متساوية الاحتمال، يكون:

$$P(\{(4,1)\,,(1,4)\}) = \frac{2}{36} = \frac{1}{18}$$

إذا:

$$P\left(X = 5 \text{ و } Y = 3\right) = \frac{1}{18}$$

من الآن نكتب، $P(X = 5\,, Y = 3)$ بدل $P\left(X = 5 \text{ و } Y = 3\right)$ بقراءة الفاصلة "و". أيضا $P(X = 6\,, Y = 0) = \frac{1}{36}$ كون فقط الحادث $(3,3)$ الذي يحقق قيم X و Y. من أجل حوادث أكثر تعقيدا، تبقى التقنية المستعملة نفسها. مثلا، $P(X = 7\,, Y \leq 4) = \frac{4}{36} = \frac{1}{9}$ كون فقط 4 عناصر أو نقاط من العينة تحقق قيم $X = 7$ و $Y \leq 4$، هي $(4,3)$، $(3,4)$، $(5,2)$ و $(2,5)$.

الشعاع العشوائي (X, Y) المعرف سابقا يسمى **شعاع عشوائي متقطع**. من أجل شعاع عشوائي متقطع، الدالة $f(x, y)$ المعرفة ب $f(x, y) = P(X = x\,, Y = y)$ يمكن أن تستعمل لحساب احتمال أي حادث معرف بدلالة (X, Y).

تعريف 2.9.:

ليكن (X, Y) شعاع عشوائي ثنائي متقطع. إذا الدالة $f(x, y)$ المعرفة من $\mathbb{R}^2$ إلى $\mathbb{R}$ ب $f(x, y) = P(X = x\,, Y = y)$ تسمى **بدالة الكتلة الاحتمالية المشتركة** ل (X, Y). من الضروري الإشارة إلى أنه كون الدالة f هي دالة الكتلة الاحتمالية المشتركة للشعاع العشوائي (X, Y) وليس لشعاع آخر فإنه سيستعمل الترميز $f_{X,Y}(x, y)$.

الجدول 1.9.: قيم دالة الكتلة الاحتمالية المشتركة $f(x, y)$

		x										
		2	3	4	5	6	7	8	9	10	11	12
y	0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
	1		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$	
	2			$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$		
	3				$\frac{1}{18}$		$\frac{1}{18}$		$\frac{1}{18}$			
	4					$\frac{1}{18}$		$\frac{1}{18}$				
	5						$\frac{1}{18}$					

دالة الكتلة الاحتمالية المشتركة ل (X,Y) تعرف تماما التوزيع الاحتمالي للشعاع العشوائي (X,Y) مثلما تعرف دالة الكتلة الاحتمالية لتوزيع متغير عشوائي واحد. من أجل (X,Y) المعرف في المثال (1.9.) بدلالة رمي زوج من زهرة النرد، يوجد 21 قيمة ممكنة ل (X,Y). قيمة $f(x,y)$ من أجل القيم 21 الممكنة هي معطاة في الجدول (1.9) أعلاه. اثنتين من هاته القيم، $f(5,3)=\frac{1}{18}$ و $f(6,0)=\frac{1}{36}$ هي محسوبة أعلاه والباقي يتحصل عليه بطريقة مماثلة. نشير إلى أن دالة الكتلة الاحتمالية المشتركة $f(x,y)$ هي معرفة من أجل جميع $(x,y)\in\mathbb{R}^2$ و ليس فقط من أجل القيم 21 المعددة في الجدول (1.9) من أجل كل القيم الأخرى غير مذكورة (x,y) لدينا $f(x,y)=P(X=x\,,Y=y)=0$.

دالة الكتلة الاحتمالية المشتركة يمكن أن تستعمل لحساب احتمال أي حادث معرف بدلالة (X,Y). لتكن A أي مجموعة جزئية من $\mathbb{R}^2$. إذا:

$$P\big((X\,,Y)\in A\big)=\sum_{(x\,,y)\in A} f(x,y).$$

كون $(X\,,Y)$ هو شعاع متقطع، $f(x,y)$ هي غير معدومة من أجل أغلب قيم (x,y). على هذا الأساس، المجموع يمكن تفسيره على أنه مجموع قابل للعد حتى وإن احتوت A على عدد غير محدد من النقاط. مثلا، لتكن $A=\big\{(x,y): x=7 \text{ و } y\le 4\big\}$. تمثل هنا مجموعة نقاط غير منتهية على نصف مستقيم في $\mathbb{R}^2$. لكن من الجدول (1.9) نلاحظ أنه فقط القيم $(x,y)\in A$ التي من أجلها $f(x,y)$ غير معدومة هي $(x,y)=(7,1)$ و $(x,y)=(7,3)$. إذا:

$$P(X=7\,,Y\le 4)=P\big((X\,,Y)\in A\big)=f(7,1)+f(7,3)=\frac{1}{18}+\frac{1}{18}=\frac{1}{9}.$$

الذي يعطي نفس القيمة المحسوبة في المثال 2.1.4 باعتبار تعريف $(X\,,Y)$ وفضاء النقاط ف S. غالبا من السهل العمل مع دالة الكتلة الاحتمالية المشتركة بدل العمل مع التعريف الأساسي.

توقعات دوال أشعة المتغيرات تحسب مثل تلك المتعلقة بمتغيرات عشوائية واحدة. لتكن $g(x,y)$ دالة قيم حقيقية معرفة من أجل جميع القيم الممكنة (x,y) لشعاع المتغير المتقطع $(X\,,Y)$. إذا $g(x,y)$ هي في حد ذاتها متغير عشوائي وقيمتها المتوقعة أو أملها يعطى ب:

$$E(g(x,y)\,)=\sum_{(x\,,y)\in\mathbb{R}^2} g(x,y)f(x,y)\;\dots\dots(9.1)$$

مثال 2.9. (**تكملة للمثال 1.9**.):

من أجل $(X\,,Y)$ التي دالة كتلتها المشتركة معطاة في الجدول (1.9)، ما هي القيمة المتوسطة المتوقعة ل XY؟ بوضع $g(x,y)=xy$، نحسب $E(XY)=E\big(g(X,Y)\big)$ من خلال حساب $xyf(x,y)$ من أجل كل 21 (x,y) قيمة في الجدول (1.9) وجمع 21 حد الناتجة. نجد:

$$E(XY)=(2)(0)\frac{1}{36}+(4)(0)\frac{1}{36}+\cdots+(8)(4)\frac{1}{18}+(7)(5)\frac{1}{18}=13\frac{11}{18}.$$

الأمل الرياضي يبقى له نفس الخصائص عندما يستبدل المتغير العشوائي X بالشعاع العشوائي (X,Y). مثلا، إذا كان لدينا الدالتين $g_1(x,y)$ و $g_2(x,y)$ بحيث a، b و c هي ثوابت، إذا:

$$E(ag_1(x,y)+bg_2(x,y)+c)=ag_1(x,y)+bg_2(x,y)+c.$$

دالة الكتلة المشتركة لأي شعاع عشوائي ثنائي $(X\,,Y)$ يجب أن تتمتع بمجموعة من الخصائص. من أجل أي (x,y)، $f(x,y)\geq 0$ كون $f(x,y)$ هي احتمال. أيضا، بما أن $(X\,,Y)$ أكيد ينتمي إلى $\mathbb{R}^2$ فإن:

$$\sum_{(x\,,y)\in\mathbb{R}^2} f(x,y)=P\big((X\,,Y)\in\mathbb{R}^2\big)=1.$$

ينتج عن ذلك أنه من أجل أي دالة غير سالبة معرفة من $\mathbb{R}^2$ إلى $\mathbb{R}$ والتي تكون غير معدومة عند أغلب الثنائيات المعدودة (x,y) ويكون مجموعها 1 هي دالة كتلة احتمالية مشتركة لشعاع عشوائي ثنائي ما. بهذه الطريقة، وبتعريف $f(x,y)$، يمكننا تعريف نموذج احتمالي ل $(X\,,Y)$ بدون الحاجة للعمل بفضاء العينة الأساسي S. المثال التالي يوضح ذلك.

مثال 3.9. (**دالة الكثافة المشتركة لزهرة نرد**):

نعرف $f(x,y)$ ب:

$$f(0,0)=f(0,1)=\frac{1}{6},$$
$$f(1,0)=f(1,1)=\frac{1}{3}$$

$f(x,y)=0$ من أجل أي قيمة أخرى ل (x,y).

$f(x,y)$ هي غير سالبة ومجموع قيمها 1، إذا $f(x,y)$ هي دالة كتلة احتمالية لشعاع عشوائي ثنائي ما $(X\,,Y)$. يمكننا استعمال $f(x,y)$ لحساب احتمالات مثل $P(X=\ Y)=f(0,0)+f(1,1)=\frac{1}{2}$. كل هذا يمكن القيام به دون الحاجة للعمل بفضاء العينة الأساسي S. لكن بالمقابل، يوجد العديد من فضاءات العينات والدوال التي تقود إلى دالة الكتلة الاحتمالية المشتركة ل $(X\,,Y)$ هذه. هذه واحدة على سبيل المثال. ليكن S فضاء العينة المتشكل من 36 حالة أو نقطة الناتجة عن تجربة رمي زهرتي نرد متزنتين. ليكن المتغير X بحيث "$X=0$" في حالة إظهار زهرة النرد الأولى وجه يحمل على الأكثر الرقم 2 و "$X=1$" في حالة إظهار زهرة النرد الأولى وجه يحمل رقم أكبر من 2. ليكن بالمقابل المتغير Y بحيث "$Y=0$" في حالة إظهار زهرة النرد الثانية وجه يحمل رقم زوجي و "$Y=1$" في حالة إظهار زهرة النرد الثانية وجه يحمل رقم فردي. يترك كتمرين لإثبات أن هذا التعريف يعطي التوزيع الاحتمالي ل $(X\,,Y)$ السابق.

حتى لو نأخذ نموذج احتمالي لشعاع عشوائي $(X\,,Y)$، قد نحسب احتمالات أو توقعات تشرك فقط متغير واحد من الشعاع العشوائي. قد نرغب مثلا بمعرفة $P(X=\ 2)$. المتغير X هو متغير عشوائي بذاته، ودالة توزيعه الاحتمالي هي ممثلة دالة كتلته الهامشية $f_X(x)=P(X=\ x)$. دالة الكتلة الهامشية ل X أو Y يمكن حسابها بسهولة من دالة الكتلة المشتركة ل $(X\,,Y)$ وفق ما تشير إليه النظرية (1.9.) التالية.

نظرية 1.9.:

ليكن $(X\,,Y)$ متغير شعاعي عشوائي ثنائي ذو دالة كتلة احتمالية مشتركة $f_{X,Y}(x,y)$. إذا دالتي الكتلة الاحتمالية الهامشية ل X و Y، $\left(f_X(x)=P(X=x)\ \text{و}\ f_Y(y)=P(Y=y)\right)$، تعطيان ب:

$$f_X(x)=\sum\nolimits_{y\in\mathbb{R}}f_{X,Y}(x,y)\ \text{و}\ f_Y(y)=\sum\nolimits_{x\in\mathbb{R}}f_{X,Y}(x,y)$$

البرهان:

سوف نبرهن النتيجة من أجل $f_X(x)$ أين يكون البرهان مماثل من أجل $f_Y(y)$.

من أجل أي $x\in\mathbb{R}$، ليكن $A_x=\{(x,y):-\infty<y<\infty\}$. إذا A_x هي الخط من المستوي حيث نقطة فاصلته هي x. إذا، من أجل $x\in\mathbb{R}$ يكون لدينا:

$$\begin{aligned}
f_X(x)&=P(X=x)\\
&=P(X=x;\,-\infty<y<\infty) && (P(-\infty<y<\infty)=1)\\
&=P\big((X;Y)\in A_x\big) && \left(A_x\ \text{تعريف}\right)\\
&=\sum_{(x,y)\in A_x}f_{X,Y}(x,y)\\
&=\sum_{(x,y)\in\mathbb{R}}f_{X,Y}(x,y)
\end{aligned}$$

مثال 4.9. (**دالة الكثافة الهامشية لتجربة رمي زهرة النرد**):

باستعمال نتيجة النظرية (1.9.)، يمكننا حساب التوزيع الهامشي ل X و Y من دالة الكتلة المشتركة المعطاة في الجدول (1.9).
لحساب دالة التوزيع الهامشية ل Y، من أجل كل قيمة ممكنة ل Y، نجمع القيم الممكنة ل X. بهذه الطريقة نتحصل على:

$$\begin{aligned}
f_Y(0)&=f_{X,Y}(2,0)+f_{X,Y}(4,0)+f_{X,Y}(6,0)\\
&\quad+f_{X,Y}(8,0)+f_{X,Y}(10,0)+f_{X,Y}(12,0)\\
&=\frac{1}{6}
\end{aligned}$$

بطريقة مماثلة نتحصل على:

$f_Y(1)=\frac{5}{18}$، $f_Y(2)=\frac{2}{9}$، $f_Y(3)=\frac{1}{6}$، $f_Y(4)=\frac{1}{9}$، $f_Y(5)=\frac{1}{18}$.

نلاحظ أن $f_Y(0)+f_Y(1)+f_Y(2)+f_Y(3)+f_Y(4)+f_Y(5)=1$ مثلما هو لازم، كون هذه الاحتمالات للقيم الستة الممكنة ل Y. دالة الكتلة الاحتمالية الهامشية ل X و Y يمكن أن تستعمل لحساب الاحتمالات أو التوقعات المتعلقة فقط ب X أو فقط ب Y. لكن لحساب الاحتمالات أو التوقعات المتعلقة ب X و Y معا، يجب استعمال دالة الكثافة الاحتمالية المشتركة.

مثال 5.9. (**احتمالات زهرة النرد**):

باستعمال دالة الكتلة الهامشية ل Y المحسوبة في المثال (4.9)، يمكننا حساب:

$$P(Y<3)=f_Y(0)+f_Y(1)+f_Y(2)=\frac{1}{6}+\frac{5}{18}+\frac{2}{9}=\frac{2}{3}.$$

أيضا:

$$E(Y^3)=0^3f_Y(0)+\cdots+5^3f_Y(5)=20\frac{11}{18}.$$

كذلك، التوزيعات الهامشية ل X و Y، المعبر عنها بدالتي الكتلة الاحتمالية الهامشية $f_X(x)$ و $f_Y(y)$، لا تفسر أو تشرح تماما التوزيع المشترك ل X و Y، بحيث يوجد العديد من التوزيعات المشتركة التي لها نفس التوزيعات الهامشية الأمر الذي يجعل من الصعب أو من المستحيل تحديد دالة التوزيع الهامشية $f_{X,Y}(x,y)$ بناءا على معرفة فقط دوال التوزيع الهامشية $f_X(x)$ و $f_Y(y)$. المثال التالي يوضح ذلك.

مثال 6.9. (**نفس التوزيعات الهامشية لكن توزيعات مشتركة مختلفة**):

نعرف دالة التوزيع المشتركة ب:

$$f(0;0)=\frac{1}{12},\quad f(1;0)=\frac{5}{12},\quad f(0;1)=f(1;1)=\frac{3}{12},$$

$f(x;y)=0$ من أجل باقي قيم x الأخرى

دالة الكتلة الهامشية ل Y هي $f_Y(0)=f(0;0)+f(1;0)=\frac{1}{2}$ و $f_Y(1)=f(0;1)+f(1;1)=\frac{1}{2}$. دالة الكتلة الهامشية ل X هي $f_X(0)=\frac{1}{3}$ و $f_X(1)=\frac{2}{2}$. نتحقق أن دالة التوزيع المشتركة هنا هي مختلفة عن تلك المعطاة في المثال (4.9.) لكن بدوال هامشية ل X و Y متماثلة. إذا كما ذكرنا، لا يمكننا تحديد دالة التوزيع المشتركة بناءا فقط على دوال التوزيع الهامشية. في الواقع دالة التوزيع المشتركة تعطينا معلومات إضافية عن توزيع الشعاع العشوائي (X,Y) التي تكون غير متوفرة في التوزيعات الهامشية.

حتى هذه النقطة ناقشنا حالة الشعاع العشوائي الثنائي المتقطع. يمكننا أيضا أخذ أو اعتبار الشعاع العشوائي الذي تكون مركباته متغيرات عشوائية مستمرة. التوزيع الاحتمالي لشعاع عشوائي مستمر يمثل بدالة الكثافة مثل حالة متغير واحد.

تعريف 3.9.:

الدالة $f(x,y)$ المعرفة من $\mathbb{R}^2$ إلى $\mathbb{R}$ تسمى **دالة الكثافة الاحتمالية المشتركة** للشعاع العشوائي الثنائي المستمر (X,Y) إذا تحقق من أجل كل $A\subset\mathbb{R}^2$:

$$P\big((\mathrm{X},\mathrm{Y})\in \mathrm{A}\big)=\iint_A f(x,y)dxdy.$$

تستعمل دالة الكثافة المشتركة مثل استعمال دالة الكثافة لمتغير واحد باستثناء أن التكامل هنا هو تكامل ثنائي بالنسبة لمجموعة تعريف في المستوي. الترميز $\iint_A$ يعني ببساطة أن أطراف التكامل هي مجموعة وأن الدالة هي متكاملة من أجل كل $(x, y) \in \mathrm{A}$. الأمل الرياضي لدالة شعاع عشوائي مستمر تعرف مثلما تعرف بالنسبة لحالة المتقطع حيث يستبدل المجموع بالتكامل ودالة الكتلة تستبدل ببدالة الكثافة. بحيث إذا كانت $g(x, y)$ دالة حقيقية، فإن الأمل الرياضي ل $g(\mathrm{X}, \mathrm{Y})$ يعطى ب:

$$E\big(g(\mathrm{X}, \mathrm{Y})\big) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(\mathrm{x}, \mathrm{y}) f(x, y)\, dx dy. \dots\dots (9.2).$$

من المهم الإشارة إلى أن دالة الكثافة الاحتمالية المشتركة هي معرفة من أجل كل $(x, y) \in \mathbb{R}^2$ حيث يمكن أن تكون معدومة من أجل مجموعة كبيرة من A إذا كان $P\big((\mathrm{X}, \mathrm{Y}) \in \mathrm{A}\big) = 0$.

دالتي الكثافة الاحتمالية الهامشية ل X و Y هي معرفة أيضا مثل حالة المتغير المتقطع، أين يحل التكامل محل المجموع، و يمكن أن تستعمل لحساب الاحتمالات أو التوقعات (الامل الرياضي) المتعلقة ب X أو Y. تعطى علاقتهما ب:

$$\begin{cases} f_X(x) = \displaystyle\int_{-\infty}^{+\infty} f(x, y)\, dy, & -\infty < y < \infty \\ f_Y(y) = \displaystyle\int_{-\infty}^{+\infty} f(x, y)\, dx, & -\infty < x < \infty \end{cases} \dots\dots\dots\dots (9.3).$$

أي دالة $f(x, y)$ تحقق $f(x, y) \geq 0$ من أجل كل $(x, y) \in \mathbb{R}^2$ و $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)\, dx dy = 1$ هي دالة كثافة احتمالية مشتركة لشعاع عشوائي مستمر ما (X, Y). كل هذه المفاهيم المتعلقة بدوال الكثافة المشتركة هي موضحة في المثالين التاليين.

مثال 7.9. (حساب الاحتمالات المشتركة):

نعرف دالة الكثافة المشتركة ب:

$$f(x, y) = \begin{cases} 6xy^2 & 0 < x < 1 \text{ و } 0 < y < 1 \\ 0 & \text{عدا ذلك} \end{cases}$$

(من الآن فصاعدا، سيفهم أن $f(x, y) = 0$ من أجل قيم (x, y) غير مذكورة في التعريف). أولا، سنتحقق أن $f(x, y)$ هي فعلا دالة كثافة احتمالية مشتركة. لدينا جليا من خلال تعريفها ومن أجل كل (x, y)، $f(x, y) \geq 0$. لحساب تكامل $f(x, y)$ من أجل قيم المستوي كلها، نشير أنه بما أن $f(x, y)$ هي معدومة باستثناء قيم المربع الوحدوي، فإن هذا التكامل هو نفسه التكامل على قيم المربع الوحدوي. إذا يكون لدينا:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)\, dx dy = \int_0^1\int_0^1 6xy^2\, dx dy = \int_0^1 [3x^2y^2\,]_0^1 dy = [y^3\,]_0^1$$

الآن، لنحسب بعض الاحتمالات مثل $P(X + Y \geq 1)$. لتكن $A = \{(x, y): x + y \geq 1\}$، يمكننا إعادة كتابة الاحتمال $P\big((X, Y) \in A\big)$. لحساب هذا الاحتمال، من التعريف (3.9.)، نكامل دالة الكثافة المشتركة على المجموعة A. لكن دالة الكثافة المشتركة تساوي 0 عدى على قيم المربع الوحدوي. إذا التكامل على A في هذه الحالة هو نفسه التكامل على قيم A التي تنتمي إلى المربع

الوحدوي. المجموعة A هي نصف المستوي في الطرف الأعلى على اليمين، والجزء A من المربع الوحدوي هي المنطقة الممثلة بالمثلث المحدود بالمستقيمات $x = 1$، $y = 1$ و $x + y = 1$. يمكننا كتابة:

$$A = \{(x, y): x + y \geq 1; 0 < x < 1; 0 < y < 1\}$$
$$= \{(x, y): x \geq 1 - y; 0 < x < 1; 0 < y < 1\}$$
$$= \{(x, y): 1 - y \leq x < 1; 0 < y < 1\}$$

هذه الأخيرة تعطينا أطراف التكامل التي نحتاجها لحساب الاحتمالات. لدينا:

$$P(X + Y \geq 1) = \iint_A f(x, y)dxdy = \int_0^1 \int_{1-y}^1 6xy^2\, dxdy = \frac{9}{10}.$$

باستعمال (.3.9)، يمكننا حساب دالتي الكثافة الهامشية ل X و Y. مثلا، لحساب $f_X(x)$، نلاحظ أنه من أجل $x \geq 1$ أو $x \leq 0$، فإن "$f(x, y) = 0$" من أجل كل قيم y. في هذه الحالة يكون لدينا:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)dy = 0.$$

من أجل $0 < x < 1$، $f(x, y)$ هي غير معدومة فقط إذا كان $0 < y < 1$. إذا من أجل $0 < x < 1$ يكون لدينا:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)dy = \int_0^1 6xy^2\, dy = [2xy^3\,]_0^1 = 2x.$$

كما أشرنا، دالة الكثافة الهامشية ل X يمكن أن تستعمل لحساب الاحتمالات المتعلقة ب X فقط. مثلا:

$$P\left(\frac{1}{2} < X < \frac{3}{4}\right) = \int_{\frac{1}{2}}^{\frac{3}{4}} 2xdx = \frac{5}{16}.$$

مثال 8.9. (**حساب الاحتمالات المشتركة –II**):

كمثال آخر لدالة الكثافة المشتركة، لتكن $f(x, y) = e^{-y}, 0 < x < y < \infty$، على الرغم من أن e^{-y} هي غير مرتبطة أو تابعة ل x، لكن $f(x, y)$ هي دالة تابعة ل x بما أن مجموعة التعريف التي من أجلها $f(x, y)$ هي غير معدومة هي تابعة ل x. سيكون هذا أكثر وضوحا باستعمال مؤشر الدالة لكتابة:

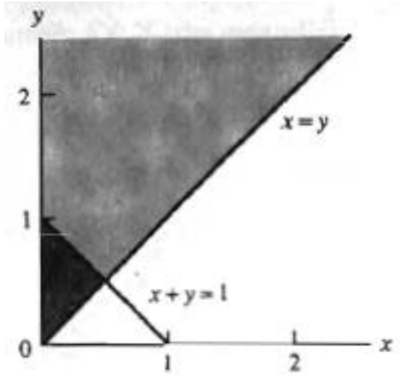
$$f(x, y) = e^{-y} I_{\{(u,v): 0<u<v<\infty\}}(x, y).$$

لحساب $P(X + Y \geq 1)$، يمكننا مكاملة دالة الكثافة المشتركة على المجال الذي تتقاطع فيه المجموعة $A = \{(x, y): x + y \geq 1\}$ والمجموعة التي من أجلها تكون $f(x, y)$ غير معدومة. الشكل 1.1.4 التالي يوضح أن مجموعة التقاطع هذه هي مجال غير محدود ذو أطراف ثلاثة معطاة ب $x = y$، $x + y = 1$ و $x = 0$. لحساب التكامل على هذا المجال يجب أن نقسمه على الأقل إلى قسمين من أجل الكتابة الملائمة لأطراف مجال التكامل.

يكون التكامل أسهل على مجموعة التقاطع بين المجموعة $b=\{(x,y): x+y<1\}$ والمجموعة التي من أجلها تكون $f(x,y)$ غير معدومة، الممثلة بالمثلث في الشكل (1.9) أدناه والمحددة أطرافه أيضا ب $x=y$ ، $x+y=1$ و $x=0$. إذا يصبح:

$$P(X+Y\geq 1)=1-P(X+Y<1)=1-\int_0^{\frac{1}{2}}\int_x^{1-x}e^{-y}dxdy$$

$$=1-\int_0^{\frac{1}{2}}\left(e^{-x}-e^{-(1-x)}\right)dx=2e^{-1/2}-e^{-1}.$$



الشكل 1.9.: مجالات المثال 8.9.

هذا يوضح أنه غالبا من القيام بالرسم البياني لمجموعة التعريف الموافقة لتحديد أطراف التكامل الملائمة للإشكاليات المماثلة.

التوزيع الاحتمالي المشترك ل (X,Y) يمكن أن يمثل بدالة الكثافة التراكمية المشتركة بدلا من دالة الكتلة أو دالة الكثافة المشتركة. دالة الكثافة التراكمية المشتركة $F(x,y)$ هي معرفة من أجل كل $(x,y)\in\mathbb{R}^2$ ب:

$$F(x,y)=P(X\leq x, Y\leq y)$$

نشير أن دالة الكثافة التراكمية المشتركة هي في الغالب غير فعالة الاستعمال في حالة شعاع عشوائي متقطع. لكن بالنسبة لشعاع عشوائي مستمر لدينا العلاقة المهمة (كذلك بالنسبة للمتغير الأحادي):

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(s,t)\,dsdt.$$

من النظرية الأساسية للحساب للمتغيرات الثنائية، هذا يستلزم:

$$\frac{\partial^2F(x,y)}{\partial x\partial y}=f(x,y).$$

عند النقاط المستمرة ل $f(x,y)$. العلاقة هاته هي مفيدة في الحالات التي يمكن فيها إيجاد صيغة ل $F(x,y)$. المشتقة الجزئية يمكن حسابها لإيجاد دالة الكثافة الاحتمالية المشتركة.

9.ب. التوزيعات الشرطية والاستقلالية:

في أغلب الأحيان عندما يتم مشاهدة متغيرين عشوائيين (X,Y) ، تكون قيمهما مرتبطة. مثلا، نفترض أنه، في تجربة اختيار عينة أشخاص عشوائية من مجتمع ما، أين X يمثل طول الشخص و Y يمثل وزن نفس الشخص. من المؤكد أنه من المحتمل أن يكون $Y > 200$ كغ لما نتكلم عن $X = 185$ سم مقارنة ب $X = 110$ سم. إذا معرفة قيمة عن X يعطينا بعض المعلومات عن قيمة Y حتى وإن كانت غير دقيقة تماما. الاحتمالات الشرطية ل Y بدلالة معرفة قيمة X يمكن أن تستعمل باستعمال التوزيع المشترك ل (X,Y). لكن أحيانا معرفة X لا تعطينا أية معلومة عن Y.

تعريف 4.9.:

ليكن (X,Y) شعاع عشوائي ثنائي متقطع ذو دالة الكتلة المشتركة $f(x,y)$ ودوال الكتلة الهامشية $f_X(x)$ و $f_Y(y)$. من أجل أي قيمة ل x بحيث $P(X=x) = f_X(x) > 0$، دالة الكتلة (الكثافة) الشرطية ل Y بدلالة $X = x$ هي دالة ل y يرمز لها $f(y/x)$ ومعرفة ب:

$$f(y/x) = P(Y = y/X = x) = \frac{f(x,y)}{f_X(x)} \dots \dots (9.4).$$

ومن أجل y بحيث يكون $P(Y=y) = f_Y(y) > 0$، دالة الكثافة الشرطية ل X بحيث $Y = y$ هي دالة ل x يرم لها $f(x/y)$ ومعرفة ب:

$$f(x/y) = P(X = x/Y = y) = \frac{f(x,y)}{f_Y(y)} \dots \dots (9.5).$$

بما أننا سمينا $f(y/x)$ دالة كثافة احتمالية، يجب أن نتحقق أن هذه الدالة بدلالة y تعرف دالة كثافة احتمالية لمتغير عشوائي. أولا، لدينا الشرط $f(y/x) \geq 0$ هو محقق بما أن $f(x,y) \geq 0$ و $f_X(x) \geq 0$. ثانيا،

$$\sum_y f(y/x) = \frac{\sum_y f(x,y)}{f_X(x)} = \frac{f_X(x)}{f_X(x)} = 1.$$

إذا $f(y/x)$ هي دالة كثافة احتمالية ويمكن أن تستعمل بالطريقة العادية لحساب الاحتمالات المتعلقة ب Y عند معرفة أن $X = x$ قد تحققت.

مثال 9.9. (**حساب الاحتمالات الشرطية**):

نعرف دالة الكتلة الاحتمالية المشتركة ل (X,Y) ب:

$$f(0;10) = f(0;20) = \frac{2}{18};\ f(1;10) = f(1;30) = \frac{3}{18};$$

$$f(1;20) = \frac{4}{18}\ \text{و}\ f(2;30) = \frac{4}{18}.$$

يمكننا أن نستعمل التعريف (4.9) لحساب دالة الكتلة الاحتمالية الشرطية ل Y بحيث X معطى من أجل جميع القيم الممكنة ل X، $x = 0,1,2$. أولا، دالة الكتلة الهامشية ل X هي:

$$f_X(0) = f(0;10) + f(0;20) = \frac{4}{18}$$

$$f_X(1) = f(1;10) + f(1;20) + f(1;30) = \frac{10}{18}$$

$$f_X(2) = f(2;30) = \frac{4}{18}.$$

من أجل $x = 0$، $f(0;y)$ هي موجبة فقط من أجل $y = 10$ و $y = 20$. إذا $f(y/0)$ هي موجبة فقط من أجل $y = 10$ و $y = 20$ ويصبح:

$$f(10/0) = \frac{f(0;10)}{f_X(0)} = \frac{\frac{2}{18}}{\frac{4}{18}} = \frac{1}{2},$$

$$f(20/0) = \frac{f(0;20)}{f_X(0)} = \frac{\frac{2}{18}}{\frac{4}{18}} = \frac{1}{2}.$$

إذا، بمعرفة أن $X = 0$، توزيع الاحتمال الشرطي ل Y هو التوزيع المتقطع الذي يعطي الاحتمال $(1/2)$ لقيمتيه: $y = 10$ و $y = 20$.

من أجل $x = 1$، $f(y/1)$ هي موجبة من أجل $y = 10$ ، $y = 20$ $y = 30$، و:

$$f(10/1) = f(30/1) = \frac{3}{18}\Big/\frac{10}{18} = \frac{3}{10},$$

$$f(20/1) = \frac{4}{18}\Big/\frac{10}{18} = \frac{4}{10}.$$

من أجل $x = 2$،

$$f(30/2) = \frac{4}{18}\Big/\frac{4}{18} = 1.$$

النتيجة الأخيرة تعكس حقيقة، تظهر أيضا من دالة الكتلة المشتركة، أنه إذا علمنا أن $X = 2$ فإنه بالضرورة يكون $Y = 30$.

احتمالات شرطية أخرى يمكن حسابها باستعمال دوال الكتلة الاحتمالية الشرطية. مثلا:

$$P(Y > 10/X = 1) = f(20/1) + f(30/1) = \frac{7}{10}$$

أو

$$P(Y > 10/X = 0) = f(20/0) = \frac{1}{2}.$$

إذا كان X و Y متغيرين مستمرين، إذا يكون $P(X = x) = 0$ من أجل كل قيمة x. لحساب الاحتمال الشرطي مثل $P(Y > 200/X = 73)$، التعريف (4.9) لا يمكن أن يستعمل كون المقام $P(X = 73)$ يكون معدوما على الرغم من أن قيمة X هي مشاهدة. لو نشاهد $X = 73$ ، هذه المعلومة تعطينا في الواقع معلومات عن Y (مثل مثال الوزن والطول في بداية الفقرة). في هذه الحالة (X و Y متغيرين مستمرين) يكون من الملائم لحساب الاحتمال الشرطي من أجل Y عند مشاهدة $X = x$ يكون باستبدال دوال الكتلة الاحتمالية بدوال الكثافة الاحتمالية.

تعريف 5.9:.

ليكن (X, Y) شعاع عشوائي ثنائي مستمر ذو دالة كثافة مشتركة $f(x, y)$ ودوال كثافة هامشية $f_X(x)$ و $f_Y(y)$. من أجل أي x بحيث $f_X(x) > 0$، دالة الكثافة الشرطية ل Y عند مشاهدة $X = x$ هي دالة ل y ، $f(y/x)$، والمعرفة ب:

$$f(y/x) = \frac{f(x, y)}{f_X(x)}.$$

أيضا، من أجل أي y بحيث $f_Y(y) > 0$، دالة الكثافة الشرطية ل X عند مشاهدة $Y = y$ هي دالة ل x ، $f(x/y)$، والمعرفة ب:

$$f(x/y) = \frac{f(x, y)}{f_Y(y)}.$$

للتأكد من كون $f(y/x)$ و $f(x/y)$ هما دالتي كثافة احتمالية، يمكن اتباع نفس خطوات التأكد تلك المحددة في التعريف (4.9) وذلك باستبدال المجاميع بالتكاملات.

إضافة إلى أهميتها في حساب الاحتمالات، دوال الكثافة أو الكتلة الشرطية يمكن أيضا أن تستعمل لحساب التوقعات. نذكر أن $f(y/x)$ كدالة ل y هي دالة كثافة أو كتلة احتمالية وتستعمل بنفس الطريقة التي استعملنا بها سابقا دوال الكثافة أو الكتلة الاحتمالية اللاشرطية. إذا كانت $g(y)$ هي دالة ل Y، إذا قيمة التوقع الشرطي ل $g(y)$ بحيث $X = x$، يرمز لها ب $E(g(y)/x)$، تعطى ب:

$$E(g(y)/x) = \sum_{y} g(y) f(y/x) \dots \dots (9.6)$$

و

$$E(g(y)/x) = \int_{-\infty}^{\infty} g(y) f(y/x) dy \dots \dots (9.7)$$

لحالتي المتغير المتقطع والمتغير المستمر على التوالي.

مثال 10.9. (**حساب دوال الكثافة الاحتمالية الشرطية**):

كما في المثال (.9.9)، ليكن الشعاع العشوائي المستمر (X, Y) ذو دالة الكثافة الاحتمالية المشتركة $f(x, y) = e^{-y}, 0 < x < y < \infty$. نرغب على سبيل المثال بحساب دالة الكثافة الاحتمالية الشرطية ل Y بحيث $X = x$. دالة الكثافة الاحتمالية الهامشية ل X تحسب على النحو التالي:

لما $x \leq 0$ لدينا $f(x, y) = 0$ من أجل قيم y، إذا $f_X(x) = 0$. في حالة $x > 0$، $f(x, y) > 0$ فقط لما $x < y$. إذا:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy = \int_{x}^{\infty} e^{-y}dy = e^{-x}.$$

حيث نلاحظ أن X له توزيع هامشي أسي.

من التعريف (.5.9)، التوزيع الشرطي ل Y بحيث $X = x$ يمكن حسابه من أجل $(0 < x)$ (بما أنه هذه القيم هي التي من أجلها $f_X(x) > 0$). من أجل x موافق لهذا يكون لدينا:

$$f(y/x) = \frac{f(x, y)}{f_X(x)} = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, \qquad \text{لما } y > x$$

و

$$f(y/x) = \frac{f(x, y)}{f_X(x)} = \frac{0}{e^{-x}} = 0, \qquad \text{لما } y > x.$$

على هذا، لما $(X = x)$، Y لديه توزيع أسي، أين x يمثل معلمة الموضع لتوزيع Y و $\beta = 1$ يمثل معلمة القياس. كذلك التوزيع الشرطي ل Y هو مختلف من أجل كل قيمة ل x. يتبع من ذلك:

$$E(Y/X = x) = \int_{x}^{\infty} ye^{-(y-x)}dy = 1 + x.$$

تباين التوزيع الاحتمالي المشخص ب $f(y/x)$ يسمى ب "**التباين الشرطي ل Y لما $X = x$**". باستعمال الترميز $V(Y/x)$

لهذا و باستعمال التعريف العادي للتباين يكون لدينا:

$$V(Y/x) = E(Y^2/x) - \left(E\left(\frac{Y}{x}\right)\right)^2 \ldots \ldots (9.8)$$

بتطبيق هذه العلاقة على مثالنا نتحصل على:

$$V(Y/x) = \int_{x}^{\infty} y^2 e^{-(y-x)}dy - \left(\int_{x}^{\infty} ye^{-(y-x)}dy\right)^2 = 1.$$

في هذا المثال، التباين الشرطي ل Y لما $X = x$ هو نفسه من أجل جميع قيم x. في حالات أخرى، قد يكون مختلف من أجل مختلف قيم x. هذا التباين الشرطي يمكن مقارنته بالتباين اللاشرطي ل Y. التوزيع الهامشي ل Y هو توزيع غاما $G(2;1)$، ذو تباين $V(Y) = 2$ حيث نلاحظ أنه بمعرفة $X = x$، فإن تشتت Y يتقلص بشكل كبير.

نستعرض الآن حالة ملموسة أين النموذج في المثال 4.2.4 يمكن استعماله. نفرض أنه لدينا مصباحين كهربائيين. مدتي فترة حياتهما هما متغيرين عشوائيين مستقلين نرمز لهما تواليا ب X و Z ولهما نفس دالة الكثافة الاحتمالية e^{-x}، $x > 0$. يشغل المصباح الأول إلى حين احتراقه ثم يشغل المصباح الثاني. نشاهد الآن X الوقت حتى يحترق المصباح الأول و $Y = X + Z$ الوقت حتى يحترق المصباح الثاني. يكون لدينا $X = x$ هي المدة حتى يحترق المصباح الأول و $Y = Z + x$ هي المدة حتى يحترق المصباح الثاني. كما في المثال (9.9.)، قيمة x تلعب دور معلمة موضع لدالة الكثافة الاحتمالية ل Y وفي هذه الحالة دالة الكثافة الاحتمالية الشرطية ل Y بحيث يعطى $X = x$ هي $f(y/x) = f_Z(y-x) = e^{-(y-x)}$، $y > x$.

التوزيع الشرطي ل Y بحيث يعطى $X = x$ يمكن أن يكون توزيع احتمالي مختلف من أجل كل قيمة ل x. من أجل هذا لدينا في الحقيقة عائلة توزيعات احتمالية ل Y من أجل كل x. عندما نرغب بوصف العائلة بأكملها، نستعمل العبارة " **التوزيع ل Y/X** ". إذا كان، على سبيل المثال، X متغير طبيعي والتوزيع الشرطي ل Y بحيث يعطى $X = x$ هو توزيع ثنائي الحد $B(x;p)$، إذا يمكننا أن نقول أن توزيع (Y/X) هو توزيع ثنائي الحد $B(X;p)$ ونكتب $(Y/X) \backsim B(X;p)$. عندما نستعمل الترميز (Y/X) أو لدينا متغير عشوائي كمعلمة لتوزيع احتمالي، نحن نصف في هذه الحالة عائلة توزيعات احتمال شرطية. توزيعات الكثافة الاحتمالية المشتركة أو توزيعات الكتلة الاحتمالية أحيانا تعرف أو تشخص بالدالة الشرطية $f(y/x)$ أو الهامشية $f_X(x)$. من التعريف نجد $f(x,y) = f(y/x)f_X(x)$.

نشير أيضا أن $E(g(Y)/x)$ هو دالة ل x. أي، من أجل كل قيمة ل x ، $E(g(Y)/x)$ هو عدد حقيقي يتحصل عليه بحساب التكامل أو المجموع المناسب. ومنه، $E(g(Y)/X)$ هو متغير عشوائي أين قيمته مرتبطة بقيمة X. إذا كان $X = x$، قيمة المتغير العشوائي $E(g(Y)/X)$ هي $E(g(Y)/x)$. من هذا، في المثال (10.9)، يمكننا كتابة $E(Y/X) = X + 1$.

في كل الأمثلة السابقة، التوزيع الشرطي ل Y حيث يعطى $X = x$ كان مختلفا من أجل القيم المختلفة ل x. في بعض الحالات، معرفة $X = x$ لا يعطينا أي معلومة إضافية حول Y عن تلك التي لدينا. هذه العلاقة المهمة بين X و Y تسمى بعلاقة "**الاستقلالية**".

تعريف 6.9.:

ليكن (X,Y) شعاع عشوائي ثنائي ذو دالة كثافة أو كتلة احتمالية $f(x,y)$ ودوال كثافة أو كتلة هامشية $f_X(x)$ و $f_Y(y)$. X و Y يسميان **متغيرات عشوائية مستقلة** إذا كان، من أجل كل $x \in \mathfrak{R}$ و $y \in \mathfrak{R}$:

$$f(x,y) = f_X(x)f_Y(y) \ldots\ldots (9.9)$$

إذا كان X و Y مستقلين، دالة الكثافة الاحتمالية الشرطية ل Y لما $X = x$ هي:

$$f(y/x) = \frac{f(x,y)}{f_X(x)} \quad \text{حسب التعريف}$$

$$= \frac{f_X(x)f_Y(y)}{f_X(x)}$$

$$= f_Y(y)$$

بغض النظر عن قيمة x. ومنه، من أجل أي $A \subset \Re$ و $x \in \Re$، $P(Y \in A/x) = \int_A f(y/x)dy = \int_A f_Y(y)dy = P(Y \in A)$. معرفة أن $X = x$ لا يعطينا أي معلومة إضافية حول Y.

التعريف (6.9.) يستعمل بطريقتين. يمكننا البدء بدالة الكثافة الاحتمالية المشتركة ثم التحقق إن كان X و Y مستقلين. للقيام بهذا يجب أن نتحقق أن العلاقة (9.9.) هي محققة أو صحيحة من أجل كل قيمة ل x و y. أو يمكننا وضع نموذج الذي من أجله يكون X و Y مستقلين. باعتبار ما يمثله X و Y يمكن أن يدل أن معرفة $X = x$ لا يعطينا أية معلومة حول Y. في هذه الحالة يمكننا تشخيص التوزيعات الهامشية ل X و Y وبعدها تشخيص أو تحديد التوزيع المشترك كحاصل جداء حسب ما هو معطى في العلاقة (9.9).

مثال 11.9. (**التحقق من الاستقلالية–I**):

ليكن الشعاع العشوائي الثنائي المتقطع (X, Y)، دالة كتلته الاحتمالية المشتركة معطاة ب:

$$f(10;1) = f(20;1) = f(20;2) = \frac{1}{10},$$

$$f(10;2) = f(10;3) = \frac{1}{5} \text{ و } f(20;3) = \frac{3}{10}.$$

دوال الكتلة الهامشية يمكن حسابها بسهولة:

$$f_X(10) = f_X(20) = \frac{1}{2};\ f_Y(1) = \frac{1}{5};\ f_Y(2) = \frac{3}{10} \text{ و } f_Y(3) = \frac{1}{2}.$$

المتغيرات العشوائية X و Y هي غير مستقلة لأن العلاقة (1.2.4) هي غير محققة من أجل جميع قيم x و y. مثلا:

$$f(10;3) = \frac{1}{5} \neq \frac{1}{2}\frac{1}{2} = f_X(10)f_Y(3).$$

نذكر أن العلاقة (9.9.) يجب أن تتحقق من أجل كل اختيار ل x و y لكي يكون X و Y مستقلين. نلاحظ أن $f(10;1) = \frac{1}{10} = \frac{1}{2}\frac{1}{5} == f_X(10)f_Y(1)$. حيث أن تحققها من أجل بعض القيم ل x و y لا يضمن أن يكون X و Y مستقلين. يجب أن تتحقق من أجل جميع القيم.

التحقق من استقلالية X و Y مباشرة من العلاقة (9.9.) يتطلب معرفة $f_X(x)$ و $f_Y(y)$. النظرية التالية تجعل التأكد من ذلك أسهل.

نظرية 2.9.:

ليكن (X, Y) شعاع عشوائي ثنائي ذو دالة كثافة أو كتلة احتمالية $f(x;y)$. يكون X و Y متغيرين عشوائيين مستقلين إذا وفقط إذا وجدت دالتين $g(x)$ و $h(x)$ بحيث من أجل كل $x \in \Re$ و $y \in \Re$ يتحقق:

$$f(x;y) = g(x)h(y).$$

البرهان:

الجزء الثاني "وفقط إذا" يبرهن بوضع أو تعريف $g(x) = f_X(x)$ و $h(y) = f_Y(y)$. لبرهنة الجزء "إذا" من أجل متغيرات عشوائية مستمرة نفرض أن $f(x;y) = g(x)h(y)$. و نعرف:

$$\int_{-\infty}^{\infty} g(x)dx = c \text{ و } \int_{-\infty}^{\infty} h(y)dy = d,$$

أين الثابتين c و d يحققين:

$$\begin{aligned} cd &= \left(\int_{-\infty}^{\infty} g(x)dx\right)\left(\int_{-\infty}^{\infty} h(y)dy\right) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy \\ &= 1. \ \left(f(x,y) \text{ هي دالة كثافة مشتركة}\right) \end{aligned}$$

أيضا، تعطى دوال الكثافة الهامشية ب:

$$f_X(x) = \int_{-\infty}^{\infty} g(x)h(y)dy = g(x)d \text{ و } f_Y(y) = \int_{-\infty}^{\infty} g(x)h(y)dx = h(y)c.$$

من هذا، وباستعمال (.9.9) يكون لدينا:

$$f(x,y) = g(x)h(y) = g(x)h(y)cd = f_X(x)f_Y(y),$$

يظهر أن X و Y مستقلين. استبدال التكامل بالمجموع يبرهن النظرية من أجل الأشعة العشوائية المتقطعة.

مثال 12.9. (**التحقق من الاستقلالية–*II***):

لتكن دالة الكثافة الاحتمالية المشتركة $f(x,y) = \frac{1}{384}x^2y^4e^{-y-(x/2)}$، $x > 0$ و $y > 0$. بتعريف:

$$g(x) = \begin{cases} x^2e^{-(x/2)} & x > 0 \\ 0 & x \le 0 \end{cases} \quad \text{و} \quad h(y) = \begin{cases} y^4e^{-y/384} & y > 0 \\ 0 & y \le 0 \end{cases} ,$$

إذا $f(x,y) = g(x)h(y)$ من أجل كل $x \in \mathfrak{R}$ و $y \in \mathfrak{R}$. من خلال النظرية (.2.9)، نستنتج أن X و Y هما متغيرين عشوائيين مستقلين. ليس علينا حساب دوال الكثافة الهامشية.

إذا كان X و Y متغيرين عشوائيين مستقلين، إذا حسب التعريف (4.9) فإنه من الواضح أن $f(x,y) > 0$ في المجموعة $\{(x,y): x \in A \text{ و } y \in B\}$، أين $A = \{x: f_X(x) > 0\}$ و $B = \{y: f_Y(y) > 0\}$. مجموعة من هذا الشكل تسمى **جداء–تقاطعي** ويرمز له غالبا ب $A \times B$. عنصر من **جداء–تقاطعي** يمكن أن يتحقق منه بأخذ قيم x و y بشكل منفصل.

إذا كانت $f(x,y)$ دالة كثافة أو كتلة مشتركة والمجموعة أين $f(x,y) > 0$ هي ليست **جداء–تقاطعي**، إذا المتغيرين X و Y ذوي دالة الكثافة المشتركة $f(x,y)$ هما غير مستقلين. في المثال (10.9)، المجموعة $0 < x < y < \infty$ هي ليست **جداء–تقاطعي**. للتحقق من العضوية في هذه المجموعة يجب أن نتحقق ليس فقط أن $0 < x < \infty$ و $0 < y < \infty$ لكن أيضا من أن $x < y$. من هذا، المتغيرات العشوائية في المثال (7.9.) هي غير مستقلة. المثال هذا يعطي مثالا لدالة كتلة احتمالية مشتركة موجبة في مجموعة ولكن ليست **جداء–تقاطعي**.

مثال 13.9. (**نموذج احتمالي مشترك**):

كمثال لاستعمال الاستقلالية لتعريف نموذج احتمالي مشترك، نأخذ الحالة التالية. طالب جامعي من جامعة ابن خلدون تيارت تم اختياره عشوائيا والمتغير X الذي يمثل عدد آباء الطالب الأحياء المسجلين. نفرض أن التوزيع الهامشي ل X هو:

$$f_X(0) = 0{,}01;\ \ f_X(1) = 0{,}09 \text{ و } f_X(2) = 0{,}90.$$

عامل متقاعد من ولاية تيارت تم اختياره عشوائيا والمتغير Y يمثل عدد آبائه الأحياء المسجلين. نفرض أن التوزيع الهامشي ل Y هو:

$$f_Y(0) = 0{,}70;\ \ f_Y(1) = 0{,}25 \text{ و } f_Y(2) = 0{,}05.$$

من المنطقي أن يكون هذين المتغيرين هما مستقلين. معرفة عدد أباء الطالب الأحياء لا يعطينا شيئا عن عدد آباء المتقاعد الأحياء. التوزيع المشترك الوحيد ل X و Y الذي يعكس هذه الاستقلالية هو المعرف ب (4.9). على سبيل المثال:

$$f(0{,}1) = f_X(0)f_Y(1) = 0{,}0025 \text{ و } f(0{,}0) = f_X(0)f_Y(0) = 0{,}0070$$

هذا التوزيع المشترك يمكن أن يستعمل لحساب احتمالات من الشكل:

$$P(X = Y) = f(0{,}0) + f(1{,}1) + f(2{,}2)$$

$$= (0{,}01)(0{,}70) + (0{,}09)(0{,}2) + (0{,}90)(0{,}05) = 0{,}0745.$$

بعض الاحتمالات والتوقعات يمكن حسابها بسهولة لما يكون X و Y مستقلين، كما تشير إلى ذلك النظرية التالية.

نظرية 3.9.:

ليكن X و Y متغيرين عشوائيين مستقلين:

أ. من أجل أي $A \subset \Re$ و $B \subset \Re$، $P(X \in A, Y \in B) = P(X \in A,)P(\,Y \in B)$، هذا كون الحادثين، $\{X \in A\}$ و $\{Y \in B\}$ هما حادثين مستقلين.

ب. لتكن $g(x)$ دالة فقط بدلالة x و $h(y)$ دالة فقط بدلالة y. إذا:

$$E\big(g(X)h(Y)\big) = E\big(g(X)\big)E\big(h(Y)\big).$$

البرهان:

من أجل متغيرات عشوائية مستمرة، الجزء (ب) يبرهن كالتالي:

$$\begin{aligned}
E\big(g(X)h(Y)\big) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f(x,y)\,dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)\,dxdy \\
&= \int_{-\infty}^{\infty} g(x)f_X(x)dx \int_{-\infty}^{\infty} h(y)f_Y(y)dy \\
&= \left(\int_{-\infty}^{\infty} g(x)f_X(x)dx\right)\left(\int_{-\infty}^{\infty} h(y)f_Y(y)dy\right) \\
&= E\big(g(X)\big)E\big(h(Y)\big).
\end{aligned}$$

تبرهن النتيجة من أجل المتغير العشوائي المتقطع باستبدال التكاملات بالمجاميع. الجزء (أ) يمكن أن يبرهن بسلسلة من الخطوات المماثلة للتي تلك أعلاه أو باتباع الطريقة التالية. لتكن $g(x)$ دالة تأشير للمجموعة A و $h(y)$ دالة تأشير للمجموعة B . إذا $g(x)h(y)$ هي دالة تأشير للمجموعة $C \subset \mathbb{R}^2$ المعرفة ب $C = \{(x,y): x \in A, y \in B\}$. باستعمال مساواة التوقع المبرهنة منذ قليل يكون لدينا:

$$\begin{aligned}
P(X \in A, Y \in B) &= P\big((X,Y) \in C\big) = E\big(g(X)h(Y)\big) \\
&= E\big(g(X)\big)E\big(h(Y)\big) = P(X \in A)P(Y \in B).
\end{aligned}$$

مثال 14.9. (**التوقعات لمتغيرات مستقلة**):

ليكن X و Y متغيرين عشوائيين أسيين $E(1)$ مستقلين. من النظرية (2.10.) لدينا:

$$P(X \geq 4, Y < 3) = P(X \geq 4)P(\,Y < 3) = e^{-4}(1 - e^{-3}).$$

بجعل $g(x) = x^2$ و $h(y) = y$ نلاحظ أن:

$$E(X^2Y) = E(X^2)E(Y) = \big(V(X) + E(X^2)\big)E(Y) = (1 + 1^2)1 = 2.$$

النتيجة التالية المتعلقة بمجموع متغيرات عشوائية متقطعة مستقلة هي نتيجة مباشرة للنظرية (3.9).

نظرية 4.9.:

ليكن X و Y متغيرين عشوائيين مستقلين ذو دالتي توليد عزم على التوالي $M_X(t)$ و $M_Y(t)$. إذا دالة توليد العزم للمتغير العشوائي $Z = X + Y$ تعطى ب:

$$M_Z(t) = M_X(t)M_Y(t).$$

البرهان:

باستعمال تعريف دالة توليد العزم و النظرية (3.10.)، يكون لدينا:

$$M_Z(t) = E(e^{tZ}) = E\left(e^{t(X+Y)}\right) = E(e^{tX})E(e^{ty}) = M_X(t)M_Y(t).$$

مثال 15.9. (**دالة توليد العزوم لمتغيرات طبيعية**):

أحيانا النظرية (4.9.) يمكن أن تستعمل بسهولة لمعرفة توزيع Z من خلال معرفة توزيع X و Y. مثلا، ليكن $X \sim N(\mu, \sigma^2)$ و $Y \sim N(\gamma, \tau^2)$ متغيرين عشوائيين مستقلين. دالتي توليد العزوم ل X و Y هما:

$M_X(t) = e^{(\mu t + \sigma^2 t^2/2)}$ و $M_Y(t) = e^{(\gamma t + \tau^2 t^2/2)}$.

من النظرية (4.9)، دالة توليد العزوم ل $Z = X + Y$ هي:

$$M_Z(t) = M_X(t)M_Y(t) = e^{\left((\mu+\gamma)t + (\sigma^2+\tau^2)t^2/2\right)}.$$

التي تمثل دالة توليد العزوم لمتغير عشوائي طبيعي ذو متوسط "$\mu + \gamma$ " وتباين " $\sigma^2 + \tau^2$". هذه النتيجة هي مهمة كفاية لكي يعبر عنها بنظرية.

نظرية 5.9.:

ليكن $X \sim N(\mu, \sigma^2)$ و $Y \sim N(\gamma, \tau^2)$ متغيرين عشوائيين طبيعيين مستقلين. إذا المتغير العشوائي "$Z = X + Y$ " له توزيع طبيعي $N(\mu + \gamma, \sigma^2 + \tau^2)$.

إذا كانت $f(x, y)$ دالة كثافة احتمالية مشتركة للشعاع العشوائي المستمر (X, Y)، النظرية (4.9.) يمكن أن لا تتحقق من أجل مجموعة A لقيم (x, y) التي من أجلها $\iint_A \, dxdy = 0$. من أجل حالة مماثلة، X و Y يبقيان يسميان متغيرين عشوائيين مستقلين. هذا يعكس حقيقة أن دالتي كثافة احتمالية، تختلفان فقط في مجموعة مثل A، تعرفان نفس التوزيع الاحتمالي ل (X, Y). لرؤية هذا، نفرض $f(x, y)$ و $f^*(x, y)$ دالتي كثافة احتمالية متساويتين عدا في المجموعة A التي من أجلها $\iint_A \, dxdy = 0$ وليكن (X, Y) لديه دالة الكثافة الاحتمالية $f(x, y)$، وليكن (X^*, Y^*) ذو دالة الكثافة الاحتمالية $f^*(x, y)$، لتكن B أي مجموعة جزئية من $\mathbb{R}^2$. إذا:

$$P\left((X, Y) \in B\right) = \iint_B f(x, y)dxdy$$

$$= \iint_{B \cap A^c} f(x, y)dxdy$$

$$= \iint_{B \cap A^c} f^*(x, y)dxdy$$

$$= \iint_B f^*(x,y)dxdy = P\big((X^*,Y^*)\ \in B\big).$$

ومنه (X,Y) و (X^*,Y^*) لهما نفس التوزيع الاحتمالي. على سبيل المثال، لتكن $f(x,y) = e^{-x-y}$، $x > 0$ و $y > 0$ هي دالة كثافة احتمالية لمتغيرين عشوائيين مستقلين أسيين وتحقق (4.9). لكن $f^*(x,y)$ التي تساوي $f(x,y)$ باستثناء أن "$f^*(x,y) = 0$" لما "x$x = y$" هي أيضا دالة كثافة احتمالية لمتغيرين عشوائيين مستقلين أسيين على الرغم من أن النظرية (5.9) هي غير صحيحة أو غير محققة من أجل المجموعة "$A = \{(x,x): x > 0\}$".

9.ت. تحويلات المتغيرات الثنائية:

في هذه الفقرة نوسع طرق إيجاد توزيع دالة متغير عشوائي إلى حالة الأشعة العشوائية ثنائية المتغيرات.

ليكن (X,Y) شعاع عشوائي ثنائي المتغيرات ذو دالة توزيع احتمالية معلومة. نعتبر الآن شعاع عشوائي ثنائي المتغيرات آخر (U,V) معرف ب $U = g_1(X,Y)$ و $V = g_2(X,Y)$ أين $g_1(x,y)$ و $g_2(x,y)$ هما دالتين مشخصتين. إذا كانت B أي مجموعة جزئية من $\mathbb{R}^2$، إذا $(U,V) \in B$ إذا وفقط إذا كان $(X,Y) \in A$ أين "$A = \{(x,y): \big(g_1(x,y), g_2(x,y)\big) \in B\}$". ومنه $P\big((U,V) \in B\big) = \big((X,Y) \in A\big)$ و التوزيع الاحتمالي ل (U,V) هو محدد بالكامل بالتوزيع الاحتمالي ل (X,Y).

إذا كان (X,Y) شعاع عشوائي ثنائي متقطع، إذا يوجد فقط مجموعة قيم معدودة التي من أجلها دالة الكتلة الاحتمالية المشتركة ل (X,Y) هي موجبة. نرمز لهذه المجموعة A. نعرف المجموعة $B = \Big\{(u,v): u = g_1(x,y)$ و $v = g_2(x,y)$ من أجل بعض $(x,y) \in \mathrm{A}\Big\}$. إذا B هي المجموعة المعدودة للقيم الممكنة للشعاع العشوائي الثنائي المتقطع (U,V). وإذا، من أجل أي $(u,v) \in B$، $A_{u,v}$ هي معرفة لتكون $\Big\{(x,y) \in A: g_1(x,y) = u$ و $g_2(x,y) = v\Big\}$، ودالة الكتلة الاحتمالية المشتركة ل (U,V)، $f_{U,V}(u,v)$، يمكن أن تحسب من دالة الكتلة الاحتمالية المشتركة ل (X,Y) من خلال:

$$f_{U,V}(u,v) = P(U = u, V = v) = P\Big((X,Y)\ \in A_{u,v}\Big) = \sum_{(x,y)\in A_{u,v}} f_{X,Y}(x,y) \ldots\ldots (9.10)$$

مثال 16.9. (**توزيع مجموع متغيرات بواسون**):

ليكن X و Y متغيرين عشوائيين بواسونين مستقلين ذو معلمتين على التوالي θ و λ. دالة الكتلة الاحتمالية المشتركة ل (X,Y) هي:

$$f_{X,Y}(x,y) = \frac{\theta^x e^{-\theta}}{x!}\frac{\lambda^y e^{-\lambda}}{y!}, \quad x = 0,1,2 \ldots, y = 0,1,2 \ldots.$$

المجموعة A هي $\{(x,y): x = 0,1,2 \dots \text{ و } y = 0,1,2 \dots\}$. الآن نعرف $U = X + Y$ و $V = Y$, يكون،

$g_1(x,y) = x + y$ و $g_2(x,y) = y$. نعرف المجموعة B، مجموعة القيم الممكنة ل (u,v). القيم الممكنة ل v هي القيم الصحيحة غير سالبة. المتغير $v = y$ لديه نفس مجموعة القيم الممكنة. من أجل قيمة معطاة ل v، $u = x + y = x + v$ يجب أن يكون صحيحا أكبر من أو يساوي v بما أن x هو عدد صحيح غير سالب. مجموعة القيم الممكنة ل (u,v) هي المعطاة ب $B = \{(u,v): v = 0,1,2 \dots \text{ و } u = v, v+1, v+2, \dots\}$. من أجل أي $(u,v) \in B$، القيمة الوحيدة (x,y) التي تحقق $x + y = u$ و $y = v$ هي $x = u - v$ و $y = v$. ومنه، في هذا المثال، $A_{u,v}$ تتمثل دائما في الثنائية $(u - v, v)$.

من (10.9) نتحصل على دالة الكتلة الاحتمالية المشتركة ل (U,V) على النحو التالي:

$$f_{U,V}(u,v) = f_{X,Y}(u - v, v) = \frac{\theta^{u-v} e^{-\theta}}{(u-v)!} \frac{\lambda^{v} e^{-\lambda}}{v!}, v = 0,1,2 \dots \text{ و } u = v, v+1, v+2, \dots$$

في هذا المثال، من الأهمية حساب دالة الكتلة الاحتمالية الهامشية ل U. من أجل أي ثابت صحيح غير سالب u، "$f_{U,V}(u,v) > 0$" فقط من أجل $v = 0,1, \dots, u$. هذا يعطي مجموعة v قيم للقيام بالجموع للحصول على دالة الكتلة الاحتمالية الهامشية ل U. نجد:

$$f_U(u) = \sum_{v=0}^{u} \frac{\theta^{u-v} e^{-\theta}}{(u-v)!} \frac{\lambda^{v} e^{-\lambda}}{v!} = e^{-(\theta+\lambda)} \sum_{v=0}^{u} \frac{\theta^{u-v}}{(u-v)!} \frac{\lambda^{v}}{v!}, u = 0,1,2 \dots$$

يمكن أن تبسط هذه بملاحظة أنه بالضرب والقسمة على $u!$ يمكننا استعمال نظرية ثنائي الحد ونتحصل على:

$$f_U(u) = \frac{e^{-(\theta+\lambda)}}{u!} \sum_{v=0}^{u} \binom{u}{v} \lambda^{v} \theta^{u-v} = \frac{e^{-(\theta+\lambda)}}{u!} (\theta + \lambda)^{u}, u = 0,1,2 \dots$$

هذه هي دالة كتلة احتمالية لمتغير عشوائي بواسوني ذو معلمة $(\theta + \lambda)$. النتيجة هي ذات أهمية كفاية لصياغتها على شكل نظرية.

نظرية 6.9.:

إذا كان $X \sim P(\theta)$ و $Y \sim P(\lambda)$ و X و Y مستقلين فإن، $X + Y \sim P(\theta + \lambda)$. إذا كان (X,Y) شعاع عشوائي مستمر ذو دالة كثافة احتمالية مشتركة $f_{X,Y}(x,y)$، إذا دالة الكثافة الاحتمالية المشتركة ل (U,V) يمكن أن يعبر عنها بدلالة $f_{X,Y}(x,y)$. مثل السابق، $A = \{(x,y): f_{X,Y}(x,y) > 0\}$ و $B = \{(u,v): u = g_1(x,y) \text{ و } v = g_2(x,y)\}$ من أجل بعض $(x,y) \in A$. دالة الكثافة الاحتمالية المشتركة $f_{U,V}(u,v)$ تكون موجبة في المجموعة B. من أجل الصيغة البسيطة لهذه النتيجة نفرض أن التحويل $u = g_1(x,y)$ و $v = g_2(x,y)$ يعرفان تحويل متقابل من A إلى B. التحويل هو متقابل بسبب تعريف المجموعة B. فرضنا أنه من أجل كل $(u,v) \in B$ يوجد عنصر وحيد فقط $(x,y) \in A$ بحيث $(u,v) = (g_1(x,y), g_2(x,y))$. من أجل تحويل متقابل، يمكننا حل المعادلات $u = g_1(x,y)$ و $v = g_2(x,y)$ من أجل x و y بدلالة u و v. نرمز إلى هذا التحويل العكسي ب $x = h_1(u,v)$ و $y = h_2(u,v)$. الدور الذي يلعبه الاشتقاق في حالة

متغير أحادي يلعبه في حالة المتغيرات الثنائية (المتعددة بصفة عامة) قيمة تسمى بالتحويل الجاكوبي أو **المحدد الجاكوبي** الذي يرمز له ب J و المعبر عن محدد مصفوفة المشتقات الجزئية على النحو التالي ب:

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial u} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial u} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v},$$

أين:

$$\frac{\partial x}{\partial u} = \frac{h_1(u,v)}{\partial u}, \qquad \frac{\partial x}{\partial v} = \frac{h_1(u,v)}{\partial v}, \qquad \frac{\partial y}{\partial u} = \frac{h_2(u,v)}{\partial u}, \qquad \frac{\partial y}{\partial v} = \frac{h_2(u,v)}{\partial v}$$

نفرض أن J هي غير معدومة من أجل أو عند B. إذا دالة الكثافة الاحتمالية المشتركة ل (U,V) هي معدومة خارج المجموعة B وفي المجموعة B هي معطاة ب:

$$f_{U,V}(u,v) = f_{X,Y}\big(h_1(u,v), h_2(u,v)\big)|J| \ldots\ldots (9.11)$$

أين $|J|$ هي القيمة المطلقة ل J. عندما نستعمل (11.9)، يكون أحيانا من الصعب تحديد المجموعة B والتحقق من أن التحويل هو متقابل بحيث يمكن التبديل إلى العلاقة (11.9). نشرح هذه العناصر من خلال الأمثلة التالية.

مثال 17.9. (**توزيع جداء متغيرات بيتا**):

ليكن $X\sim B(\alpha,\beta)$ و $Y\sim B(\alpha+\beta,\gamma)$ متغيرين عشوائيين مستقلين. دالة الكثافة الاحتمالية المشتركة ل (X,Y) هي:

$$f_{X,Y}(x,y) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}\frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha+\beta)\Gamma(\gamma)}y^{\alpha+\beta-1}(1-y)^{\gamma-1},$$
$$0<x<1, \qquad 0<y<1.$$

وليكن التحويل $U = XY$ و $V = X$. مجموعة القيم الممكنة ل V هي $0 < v < 1$ بما أن $V = X$. من أجل قيمة محددة أو ثابتة $V = v$. U يجب أن تكون بين 0 و v بما أن $X = V = v$ و Y هي محصورة بين 0 و 1. ومنه، هذا التحويل يربط المجموعة A بالمجموعة $B = \{(u,v): 0 < u < v < 1\}$.من أجل أي $(u,v) \in B$، المعادلة $u = xy$ و $v = x$ يمكن أن تحل فقط من أجل $x = h_1(u,v) = v$ و $y = h_2(u,v) = u/v$. نشير أنه باعتبار التحويل معرف على كل $\mathbb{R}^2$، هذا التحويل ليس تحويل متقابل. أي نقطة $(0,y)$ تقابل النقطة $(0,0)$. لكن كدالة معرفة فقط على A، هي تحويل متقابل. قيمة المحدد الجاكوبي تعطى ب:

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial u} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ \frac{1}{v} & -\frac{u}{v^2} \end{vmatrix} = -\frac{1}{v}.$$

ومنه و من خلال (10.9) نتحصل على دالة الكثافة الاحتمالية المشتركة:

$$f_{U,V}(u,v)=\frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)}v^{\alpha-1}(1-v)^{\beta-1}\left(\frac{u}{v}\right)^{\alpha+\beta-1}(1-\frac{u}{v})^{\gamma-1}\frac{1}{v},$$
$$0<u<v<1\ \ldots\ldots(9.12).$$

التوزيع الهامشي ل $X = V$ هو، بالطبع، توزيع $B(\alpha,\beta)$. لكن توزيع U هو أيضا توزيع بيتا:

$$f_U(u)=\int_u^1 f_{U,V}(u,v)dv$$
$$=\frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)}u^{\alpha-1}\int_u^1\left(\frac{u}{v}-u\right)^{\beta-1}(1-\frac{u}{v})^{\gamma-1}\left(\frac{u}{v^2}\right)dv.$$

العلاقة (12.9) استعملت لكن بإعادة ترتيب بعض عناصرها. الآن نقوم بتغيير أحادي للمتغير $y=(u/v-u)/(1-u)$ بحيث $dy=-u/[v^2(1-u)]dv$ لنتحصل على:

$$f_U(u)=\frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)}u^{\alpha-1}(1-u)^{\beta+\gamma-1}\int_0^1 y^{\beta-1}(1-y)^{\gamma-1}dy,$$
$$=\frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)}u^{\alpha-1}(1-u)^{\beta+\gamma-1}\frac{\Gamma(\beta)\Gamma(\gamma)}{\Gamma(\beta+\gamma)}$$
$$=\frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta+\gamma)}u^{\alpha-1}(1-u)^{\beta+\gamma-1},\qquad 0<u<1.$$

للحصول على المساواة الثانية نذكر أن التكامل هو نواة لدالة كثافة احتمالية لبيتا حيث نستنتج أن التوزيع الهامشي ل U هو توزيع $B(\alpha,\beta+\gamma)$.

مثال 18.9. (**مجموع وفرق متغيرات طبيعية**):

ليكن X و Y متغيرين عشوائيين طبيعيين معياريين مستقلين وليكن التحويل $U=X+Y$ و $V=X-Y$. حسب الترميز المستعمل أعلاه، $U=g_1(X,Y)$ أين $g_1(x,y)=x+y$ و $V=g_2(X,Y)$ أين $g_2(x,y)=x-y$. دالة الكثافة الاحتمالية المشتركة ل X و Y هي بالطبع، $f_{X,Y}(x,y)=(2\pi)^{-1}exp(-x^2/2)exp(-y^2/2),-\infty<x<\infty,-\infty<y<\infty$. إذا المجموعة $A=\mathbb{R}^2$. لتحديد المجموعة B التي من أجلها $f_{U,V}(u,v)$ تكون موجبة، يجب أن نحدد كل القيم التي من أجله:

$$u=x+y\quad و\quad v=x-y\ \ldots\ldots\ldots(9.13)$$

التي تأخذ (x,y) مرتبة على المجموعة $A=\mathbb{R}^2$. لكن يمكننا اعتبار u أي عدد و v أي عدد وحل فقط المعادلات (13.9) من أجل x و y للحصول على:

$$x=h_1(u,v)=\frac{u+v}{2}\quad و\quad y=h_2(u,v)=\frac{u-v}{2}\ \ldots\ldots\ldots(9.14).$$

هذا يظهر أمرين. من أجل أي $(x,y) \in A$ يوجد $(u,v) \in \mathbb{R}^2$ (معرفة ب (14.9)) بحيث $u = x + y$ و $v = x - y$.
إذا، B، مجموعة كل القيم الممكنة (u,v) هي $\mathbb{R}^2$. بما أن حل (13.9) هو وحيد، هذا يظهر أيضا أن التحويل الذي لدينا هو تحويل متقابل. فقط (x,y) المعطى في العلاقة (14.9) يعطي (أو يسمح) $u = x + y$ و $v = x - y$. من العلاقة (14.9) المشتقات الجزئية ل x و y هي سهلة الحساب حيث نتحصل على:

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial u} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\ \dfrac{1}{2} & -\dfrac{1}{2} \end{vmatrix} = -\frac{1}{2}.$$

باستبدال عبارات (14.9) من أجل x و y في $f_{X,Y}(x,y)$ واستعمال $|J| = 1$، نتحصل على دالة الكثافة الاحتمالية المشتركة ل (U,V) من (12.9):

$$f_{U,V}(u,v) = f_{X,Y}\big(h_1(u,v), h_2(u,v)\big)|J| = \frac{1}{2\pi} e^{-((u+v)/2)^2/2} e^{-((u-v)/2)^2/2} \frac{1}{2}$$

من أجل $-\infty < u < \infty$ و $-\infty < v < \infty$. بتبسيط جداء الأسين نلاحظ أنه يختزل طرف uv. بعد ذلك وبعد تبسيطات وترتيبات نتحصل على:

$$f_{U,V}(u,v) = \left(\frac{1}{\sqrt{2\pi}} e^{-u^2/4}\right)\left(\frac{1}{\sqrt{2\pi}} e^{-v^2/4}\right).$$

دالة الكثافة الاحتمالية المشتركة فككت إلى جداء دالة بدلالة u ودالة بدلالة v. النظرية (6.9)، U و V هما مستقلين. من خلال النظرية (6.9)، التوزيع الهامشي ل $U = X + Y$ هو $N(0,2)$. بالتوازي، النظرية (6.9.) يمكن أن تستعمل لإيجاد أن توزيع V هو أيضا $N(0,2)$. أهمية النتيجة، أن مجموع وفرق متغيرات عشوائية طبيعية مستقلة هي متغيرات عشوائية طبيعية، و (14.9.) تعطيانا التوزيعات الهامشية ل U و V. لكن تحليل أعمق يتطلب هنا لتبيين أن U و V مستقلين.

في المثال (18.9.)، وجدنا أن U و V هما متغيرين عشوائيين مستقلين. توجد حالة بسيطة، لكن مهمة، أين متغيرين جديدين U و V، معرفين بدلالة المتغيرين الأصليين X و Y، هما مستقلين. النظرية (7.9) التالية تصف ذلك.

نظرية 7.9.:

ليكن X و Y متغيرين عشوائيين مستقلين. ليكن، $g(x)$ دالة بدلالة فقط x و $h(y)$ دالة فقط بدلالة y. إذا المتغيرات العشوائية $U = g(X)$ و $V = h(Y)$ هي مستقلة.

البرهان:

سوف نبرهن النظرية بافتراض أن U و V هي متغيرات عشوائية مستمرة. من أجل أي $u \in \mathbb{R}$ و $v \in \mathbb{R}$، نعرف:

$A_u = \{x: g(x) \le u\}$ و $B_v = \{y: h(y) \le v\}$.

إذا دالة الكثافة التراكمية المشتركة ل (U,V) هي:

$F_{U,V}(u,v) = P(U \le u, V \le v)$ (تعريف دالة الكثافة التراكمية)

$$= P(X \in A_u, Y \in B_v)$$ (تعريف U و V)

$$= P(X \in A_u)P(Y \in B_v)$$ (النظرية 3.9)

دالة الكثافة الاحتمالية المشتركة ل (U,V) هي:

$$F_{U,V}(u,v) = \frac{\partial^2}{\partial u \partial v} F_{U,V}(u,v)$$
$$= \left(\frac{d}{du} P(X \in A_u)\right)\left(\frac{d}{dv} P(Y \in B_v)\right),$$

أين، مثلما يشير إليه الترميز، المعامل الأول هو دالة فقط ل u والمعامل الثاني هو دالة فقط ل v. ومنه، حسب النظرية (7.9)، U و V هما مستقلين.

من الممكن أن نهتم بدالة وحيدة، لتكن $U = g_1(X,Y)$. في حالة مماثلة، هذه الطريقة يمكن أن تستعمل لإيجاد توزيع U. إذا كانت دالة أخرى ملائمة، $V = g_2(X,Y)$، يمكن اختيارها إذا نتيجة التحويل من (X,Y) إلى (U,V) هي متطابقة على A، إذا دالة الكثافة الاحتمالية المشتركة ل (U,V) يمكن أن تشتق (تستنتج) باستعمال العلاقة (11.9) و دالة الكثافة الاحتمالية الهامشية ل U يمكن أن يتحصل عليها من دالة الكثافة الاحتمالية المشتركة. في المثال التالي، ربما نهتم فقط ب $U = XY$. يمكننا اختيار لتعريف $V = X$، معترفين أن التحويل الناتج هو تحويل متطابق على A. إذا يمكننا مباشرة، كما في المثال، للحصول على دالة الكثافة الاحتمالية الهامشية ل U. لكن اختيارات أخرى، مثل $V = Y$، يمكن أن تعمل أيضا.

بالطبع، في حالات عديدة، التحويل هو غير متطابق. ليكن $A = \{(x,y): f_{X,Y}(x,y) > 0\}$. نفرض أن $A_0, A_1, \dots\dots, A_k$. تشكل تجزئة ل A مع الخصائص التالية. المجموعة A_0، التي يمكن أن تكون فارغة، والتي تحقق $P\big((X,Y) \in A_0\big) = 0$. التحويل $U = g_1(X,Y)$ و $V = g_2(X,Y)$ هو تحويل متقابل من A_i إلى B من أجل كل $i = 1,2, \dots\dots, k$. إذا من أجل كل i، يمكن إيجاد الدوال العكسية من B إلى A_i. نرمز للعكس i ب $x = h_{1i}(u,v)$ و $y = h_{2i}(u,v)$, هذا العكس i يعطي، من أجل $(u,v) \in B$، الثنائي الوحيد $(x,y) \in A_i$ بحيث $(u,v) = \big(g_1(x,y), g_2(x,y)\big)$. لتكن J_i ترمز إلى قيمة المحدد الجاكوبي المحسوبة من العكس i. بافتراض أن المحددات الجاكوبية هي معرفة (موجودة) على B، لدينا التمثيل التالي لدالة الكثافة الاحتمالية المشتركة، $f_{U,V}(u,v)$:

$$f_{U,V}(u,v) = \sum_{i=1}^{k} f_{X,Y}\big(h_{1i}(u,v), h_{2i}(u,v)\big)|J_i| \dots\dots (9.15).$$

مثال 19.9. **(توزيع نسبة متغيرات عشوائية طبيعية)**:

ليكن X و Y متغيرين عشوائيين طبيعيين معياريين مستقلين $N(0,1)$. ليكن التحويل $U = X/Y$ و $V = |Y|$. (U و V يمكن أن يعرفا بأي قيمة، مثلا $(1,1)$، إذا كان $Y = 0$ كون $P(Y = 0) = 0$). هذا التحويل هو غير تقابلي كون النقطتين (x,y) و $(-x,-y)$ لهما نفس الصورة أو النقطة (u,v)، لكن بفرض قيود كقيم موجبة أو سالبة على y يصبح التحويل متقابل. ليكن:

$A_1 = \{(x,y): y > 0\}$ ، $A_2 = \{(x,y): y < 0\}$ و $A_0 = \{(x,y): y = 0\}$

A_0، A_1 و A_2 تشكل تجزئة من $A = \mathbb{R}^2$ و $P\big((X,Y) \in A_0\big) = P(Y = 0) = 0$. من أجل كل A_1 أو A_2، إذا كان $(x,y) \in A_i$، $v = |y| > 0$، ومن أجل قيمة محددة $v = |y|$، $u = x/y$ يمكن ان تأخذ أي قيمة حقيقية كون x يمكن ان تأخذ أي قيمة حقيقية. إذا، $B = \{(u,v): v > 0\}$ هي صورة كل من A_1 و A_2 حسب التحويل. كذلك، التحويل العكسي من B نحو A_1 ومن B نحو A_2 يعطى ب $x = h_{11}(u,v) = uv$، $y = h_{21}(u,v) = v$ و $x = h_{12}(u,v) = -uv$، $y = h_{22}(u,v) = -v$. نشير أنه العكس الأول يعطي قيم موجبة ل y والثاني قيم سالبة ل y. المحددين الجاكوبيين من العكسيين هما $J_1 = J_2 = v$. باستعمال:

$$f_{X,Y}(x,y) = \frac{1}{2\pi} e^{-x^2/2} e^{-y^2/2},$$

من (15.9) نتحصل على:

$$\begin{aligned} f_{U,V}(u,v) &= \frac{1}{2\pi} e^{-(uv)^2/2} e^{-v^2/2} |v| + \frac{1}{2\pi} e^{-(-uv)^2/2} e^{-(-v)^2/2} |v| \\ &= \frac{v}{\pi} e^{-(u^2+1)v^2/2}, \qquad -\infty < u < \infty,\ \ 0 < v < \infty. \end{aligned}$$

من هذه دالة الكثافة الاحتمالية الهامشية ل U يمكن أن تحسب لتعطى ب:

$$\begin{aligned} f_U(u) &= \int_0^\infty \frac{v}{\pi} e^{-(u^2+1)v^2/2} dv \\ &= \frac{1}{2\pi} \int_0^\infty e^{-(u^2+1)z/2} dz \quad (z = v^2) \\ &= \frac{1}{2\pi} \frac{2}{(u^2+1)} \\ &= \frac{2}{\pi(u^2+1)}, \qquad -\infty < u < \infty. \end{aligned}$$

إذا رأينا أن قسمة (أو النسبة بين) متغيرين عشوائيين طبيعيين معياريين مستقلين هو متغير عشوائي لكوشي. (انظر التمرين 28.10. من أجل علاقات أكثر بين متغيرات طبيعية ومتغيرات كوشي).

10. تمارين:

1.10. نقطة عشوائية (X,Y) هي موزعة بانتظام على المربع ذو القمم $(1,1)$، $(1,-1)$، $(-1,1)$ و $(-1,-1)$. حيث دالة الكثافة الاحتمالية المشتركة هي $f(x,y)=\frac{1}{4}$ في المربع. حدد احتمالات الحوادث التالية:

أ. $X^2+Y^2<1$

ب. $2X-Y>0$

ت. $|X+Y|<2$

2.10. برهن الخصائص التالية للأمل الرياضي لمتغيرات ثنائية. من أجل متغيرين عشوائيين X و Y، الدالتين $g_1(x,y)$ ، $g_2(x,y)$ و الثوابت a، b و c:

أ. $E(ag_1(x,y)+bg_2(x,y)+c)=a+E\big(g_1(x,y)\big)+bE\big(g_2(x,y)\big)+c.$،

ب. إذا كان $g_1(x,y)\geq 0$ فإن $E\big(g_1(x,y)\big)\geq 0$،

ت. إذا كان $g_1(x,y)\geq g_2(x,y)$ فإن $E\big(g_1(x,y)\big)\geq E\big(g_2(x,y)\big)$،

ث. إذا كان $a\leq g_1(x,y)\leq b$ فإن $a\leq E\big(g_1(x,y)\big)\leq b$.

3.10. باستعمال التعريف 1.1.4، بين أن الشعاع العشوائي (X,Y) المعرف في نهاية المثال 5.1.4 لديه دالة الكتلة الاحتمالية المعطاة في هذا المثال.

4.10. دالة كثافة احتمالية معرفة ب:

$$f(x,y)=\begin{cases} C(x+2y) & \text{لما } 0<x<2 \text{ و } 0<y<1 \\ 0 & \text{عدا ذلك} \end{cases}$$

أ. أوجد قيمة C،

ب. أوجد التوزيع الهامشي ل X،

ت. أوجد دالة الكثافة الاحتمالية المشتركة ل X و Y،

ث. أوجد دالة الكثافة الاحتمالية للمتغير العشوائي $Z=9/(X+1)^2$.

5.10.

أ. أوجد $P\big(X>\sqrt{Y}\big)$ إذا كان X و Y لديهما دالة التوزيع المشتركة:

$$f(x,y)=x+y,\ \ 0\leq x\leq 1,\ \ 0\leq y\leq 1.$$

ب. أوجد $P(X^2<Y<X)$ إذا كان X و Y لديهما دالة التوزيع المشتركة:

$$f(x,y)=2x,\ \ 0\leq x\leq 1,\ \ 0\leq y\leq 1.$$

6.10. اتفق A و B أن يلتقيا في مكان محدد بين الواحدة (1) و الثانية (2) زوالا. نفرض أنهما أتيا إلى مكان الموعد عشوائيا وباستقلالية خلال ساعة الموعد. أوجد توزيع طول الفترة أو الزمن الذي ينتظره A لوصول B. (إذا قدم B قبل A، نعرف زمن انتظار A ب 0).

7.10. امرأة عاملة تذهب إلى العمل بين 8 و 8:30 صباحا وتستغرق بين 40 و 50 دقيقة للوصول إليه. ليكن المتغير العشوائي X الذي يمثل وقت ذهابها، والمتغير العشوائي Y الذي يمثل مدة ذهابها. بافتراض أن هذين المتغيرين مستقلين وتوزيعهما أحادي الشكل، أوجد احتمال أن تصل هاته المرأة قبل 9 صباحا.

8.10. برهن أنه إذا حققت دالة الكثافة المشتركة ل X و Y:

$$F_{X,Y}(x,y) = F_X(x)F_Y(y),$$

فإنه من أجل أي زوج من المجالات (a,b) و (c,d) فإنه:

$$P(a \le X \le b, c \le Y \le d) = P(a \le X \le b)P(c \le Y \le d).$$

9.10. المتغير الثنائي لديه التوزيع:

		X		
		1	2	3
Y	2	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
	3	$\frac{1}{6}$	0	$\frac{1}{6}$
	4	0	$\frac{1}{3}$	0

أ. بين أن X و Y مرتبطين.

ب. أعط جدول احتمالات من أجل المتغيرين العشوائيين U و V الذين لهما نفس التوزيعات الهامشية ك X و Y لكن مستقلين.

10.10. لكن U : عدد المحاولات اللازمة للحصول على أول وجه و V : عدد المحاولات اللزمة للحصول على وجهين في تجربة رمي قطعة نقدية متزنة. هل U و V متغيرين عشوائيين مستقلين؟

11.10. إذا تم تقطيع قطعة خشبية عشوائيا إلى قطع، ما هو احتمال أن يكون القطع بطريقة يجعل بالإمكان وضع القطع معا على شكل مثلث؟ (انظر Gardner 1961 من أجل أكثر تفاصيل حول هذه الإشكالية).

12.10. لكن X و Y متغيرين عشوائيين ذو متوسطين منهيين.

بين أن:

$$\min_{g(x)} E\big(Y - g(X)\big)^2 = E\big(Y - E(Y/X)\big)^2,$$

13.10. ليكن X و Y متغيرين عشوائيين طبيعيين معياريين مستقلين $N(0,1)$.

أ. أوجد $P(X^2 + Y^2 < 1)$.

ب. أوجد $P(X^2 < 1)$ بعد التأكد من أن $X^2 \sim \chi_1^2$.

14.10. ليكن $X \sim P(\theta)$ و $Y \sim P(\lambda)$، متغيرين مستقلين. تم التبيين في النظرية (2.3.4) أن توزيع $X+Y$ هو $P(\theta+\lambda)$. بين أن توزيع $X/X+Y$ هو توزيع ثنائي الحد مع $p=\theta/(\theta+\lambda)$. ما هو توزيع $Y/X+Y$؟

15.10. ليكن X و Y متغيرين عشوائيين مستقلين لهما نفس التوزيع الهندسي.

أ. بين أن U و V هما مستقلين، أين U و V هما معرفين ب:

$U=min(X,Y)$ و $V=X-Y$.

ب. أوجد توزيع $Z=X/(X+Y)$، بحيث $Z=0$ لما $X+Y=0$.

ت. أوجد دالة الكثافة الاحتمالية المشتركة ل X و $X+Y$.

16.10. ليكن X متغير عشوائي أسي $E(1)$ ونعرف Y على أنه الجزء الصحيح ل $X+1$، بحيث:

$Y=i+1$ إذا وفقط إذا كان $i\le X<i+1$، $i=0,1,2,\dots$

أ. أوجد توزيع Y . م بماذا يعرف توزيعه؟

ب. أوجد التوزيع الشرطي ل $X-1$ لما $Y\ge 5$.

17.10. لما تعطى خاصية $g(x)$ ب:

$$\int_0^\infty g(x)dx=1$$

بين أن :

$$f(x,y)=\frac{2g\left(\sqrt{x^2+y^2}\right)}{\pi\sqrt{x^2+y^2}},\quad x,y>0,$$

هي دالة كثافة احتمالية.

18.10.

أ. ليكن X_1 و X_2 متغيرين عشوائيين طبيعيين معياريين $N(0,1)$. أوجد دالة الكثافة الاحتمالية ل $(X_1-X_2)^2/2$.

ب. إذا كان $X_i, i=1,2$ متغيري غاما عشوائيين مستقلين $G(\alpha_i,1)$، أوجد التوزيعات الهامشية ل $X_1/(X_1+X_2)$ و $X_2/(X_1+X_2)$.

19.10. ليكن X_1 و X_2 متغيرين عشوائيين طبيعيين $N(0,\sigma^2)$.

أ. أوجد التوزيع المشترك ل Y_1 و Y_2 أين:

$Y_2=X_1^2+X_2^2$ و $Y_1=\frac{X_1}{\sqrt{Y_1}}$.

ب. بين أن Y_1 و Y_2 هما مستقلين، وفسر هذه النتيجة هندسيا.

20.10. تولد نقطة عشوائيا في المعلم وفقا للمخطط أو الشكل التالي. يتم اختيار قطر R، أين $R^2 \backsim \chi_2^2$. بطريقة مستقلة يتم اختيار زاوية θ، أين $(0,2\pi) \backsim$ θ. أوجد التوزيع المشترك ل $X = R\cos\theta$ و $Y = R\sin\theta$.

21.10. ليكن (X,Y) شعاع عشوائي ثنائي دالة كثافته المشتركة $f(x,y)$. ليكن $U = aX + b$ و $V = cY + d$، اين a، b، c و d هي ثوابت مع $a > 0$ و $c > 0$. بين أن دالة الكثافة المشتركة ل (U,V) هي:

$$f_{U,V}(u,v) = \frac{1}{ac} f\left(\frac{u-b}{a}, \frac{v-d}{c}\right).$$

22.10. من أجل X و Y المعطيين في المثال (17.9)، أوجد توزيع XY باستعمال التحويلين المعطيين (أ) و (ب) و أعط تكامل V:

أ. $U = XY$ ، $V = X$.

ب. $U = XY$ ، $V = X/Y$.

23.10. ليكن X و Y متغيرين عشوائيين مستقلين مع $X \backsim G(r,1)$ و $Y \backsim G(s,1)$. بين أن $Z_1 = X + Y$ و $Z_2 = X/(X+Y)$ مستقلين وأوجد توزيع كل منهما. (Z_1 غاما و Z_2 بيتا).

24.10. استعمل تقنيات الفقرة (3.4) لإيجاد (استخلاص) التوزيع المشترك ل (X,Y) من التوزيع المشترك ل (X,Z) في المثالين (8.5.4) و (9.5.4).

25.10. X و Y هما متغيرين عشوائيين مستقلين مع $X \backsim E(\lambda)$ و $Y \backsim E(\mu)$. من الصعب أو من المستحيل مشاهدة ملاحظات مباشرة ل X و Y. لكن بالمقابل نشاهد المتغيرين العشوائيين Z و W، أين:

$$Z = min\{X,Y\} \text{ و } W = \begin{cases} 1 & si\ Z = X \\ 0 & si\ Z = Y \end{cases}$$

(هذه الحالة تظهر خاصة في التجارب الطبية. المتغيرين X و Y هما متغيرا الفحص).

أ. أوجد التوزيع المشترك ل Z و W.

ب. بين أن Z و W مستقلين. (تنبيه: بين أن $P(Z \le z/W = i) = P(Z \le z)$ من أجل $i = 0$ أو 1.)

26.10. ليكن $X \backsim N(\mu,\sigma^2)$ و $Y \backsim N(\gamma,\sigma^2)$. نفرض أن X و Y مستقلين. نعرف $U = X + Y$ و $V = X - Y$. بين أن U و V هما متغيرين طبيعيين مستقلين. أوجد توزيع كل واحد منهما.

27.10. ليكن X و Y متغيرين عشوائيين طبيعيين معياريين مستقلين.

أ. بين أن $X/(X+Y)$ له توزيع كوشي.

ب. أوجد توزيع $X/|Y|$.

ت. هل جواب الجزء (ب) مفاجئ؟ هل بإمكانك صياغة نظرية عامة؟

28.10. لاحظ جونس (Jones) (1999) لتوزيع الدالتين X و Y اين $X = R\cos\theta$ و $Y = R\sin\theta$ بحيث $\theta \backsim U(0,2\pi)$ و R متغير عشوائي موجب. بين هنا حالتين من بين تلك التي لاحظها:

أ. بين أن X/Y له توزيع كوشي.

ب. بين أن توزيع $(2XY)/\sqrt{X^2+Y^2}$ هو نفس توزيع X. خصص هذه النتيجة لمتغيرات عشوائية $N(0,\sigma^2)$.

29.10. نفرض أن توزيع Y شرطيا لما $X=x$ هو $N(x,x^2)$ والتوزيع الهامشي ل X هو $U(0,1)$.

أ. أوجد $E(Y)$، $V(Y)$ و $COV(X,Y)$.

ب. برهن أن Y/X و X مستقلين.

30.10. نفرض أن المتغير العشوائي Y له توزيع ثنائي الحد $B(n,X)$، أين n هو عدد ثابت معطى و X هو متغير أحادي الشكل المستمر $U(0,1)$.

أ. أوجد $E(Y)$ و $V(Y)$.

ب. أوجد التوزيع المشترك ل X و Y.

ت. أوجد التوزيع الهامشي ل Y.

المراجـــــع

1. Anderson, T.W., An introduction to multivariate Statistical Analysis, Wiley (1984).
2. Anne-Marie DUSSAIX, Jean-Pierre INDJEHAGOPIAN, Méthodes statistique appliquées à la gestion, Les éditions d'organisations, Paris, 1981.
3. Bernard Grais, Méthodes statistique, Dunod, Paris,
4. Calot G., Cours de calcul de probabilité, Dunod, Paris, 1971.
5. Calot G., Cours de statistique déscriptive, Dunod, Paris, 1973
6. Célyne Laliberté, Probabilités et statistiques,ERPI,
7. Chibat Ahmed, Notions sur le calcul des probabilités, Imprimerie TopColors, Constantine.
8. Corina Reischer, Raymond Leblanc, Bruno Rémillard, Denis Larocque, Théorie des probabilités, Problèmes et solutions, Presse de l'université du québec , 2002.
9. Donald H.Sanders, François Allard, Les Statistique, McGraw-Hill, Canada, 1980.
10. Fourgeaud H., Fuchs A., Statistique, Dunod, Paris, 1967.
11. George Casella, Roger L. Berger « Statistical Inference », 2nd Edition, DUXBURY ADVANCED SERIES, 2002.
12. . Gilbert SAPORTA : Probabilités− Analyse des données et statistique, Technip Éditions.
13. Guitton H., Statistique, Dalloz, Paris, 1971.
14. Ludovic Lebart, Alain Morineau, Marie Piron, Statistique exploratoire multidimensionnelle, Dunod, Paris, 2000.
15. Malinvaud E., Méthodes statistiques de l'économétrie, Donod, Paris, 1964.
16. Morice E., Chartier F., Méthode statistique, Imprimerie nationale, Paris, 1954.
17. Monfort A., Cours de probabilités, Economica, Paris, 1980.
18. Monfort A., Cours de statistique mathématique, Economica, Paris, 1982.
19. Renée Veysseyre, Aide-mémoire Statistique et probabilité pour l'ingénieur, Dunod, Paris, 2006.
20. Redjal Kaci, Cours de probabilités, OPU, Algerie, 1995
21. Redjal Kaci, Probabilités exercices corrigés avec rappels de cours, K R, éditions, Algerie
22. Renyi A., Calcul des probabilités, Dunod, Paris, 1966.
23. Saporta G., Théorie et méthodes de la statistique, Technip, Paris, 1978.
24. Saporta G., Probabilités, analyse des données et statistique, Technip, Paris, 1990.
25. Tassi P., Méthodes statistiques, Economica, paris, 1985.
26. أ. ستي حميد، الإحصاء 03 محاضرات وتمارين، Alpha Edition، 2022.
27. أ. ستي حميد، الإحصاء 03 مسائل وتمارين محلولة، Alpha Edition، 2022.