الفصل السابع

التوزيعات الثنائية

BIVARIATE DISTRIBUTIONS

1-7 مقدمة Introduction

في دراستنا للتوزيعات ذات المتغيرات العشوائية في الفصول السابقة ،درسنا التوزيعات ذات المتغير العشوائي الواحد ، حيث أن المشاهدات تأخذ متغير واحد.وفي حالات كثيرة فان المشاهدات تتطلب متغيرين أو أكثر ، لذلك سوف نهتم في هذا الفصل بدراسة وتقييم العلاقات الإحصائية لهذه المتغيرات.فمثلاً في مسألة المتغيرات الثلاثة نأمل أن نتعرف أيهما من مشاهدات المتغير العشوائي الأول تزداد تبعاً لزيادة مشاهدات المتغير العشوائي الثاني ، ولكن مشاهدات المتغير العشوائي الثالث غير مرتبطة بالتغير الذي يحصل في المتغيرين السابقين .[2]

هذه المتغيرات تكون أما كمية مستمرة(Quantitative Continuous) أو نوعية متقطعة(Qualitative Discrete) أو تكون مختلطة(مستمرة ومتقطعة) . إن عدد الاحتمالات الممكنة يزداد بزيادة عدد المتغيرات العشوائية ومن الأمثلة على ذلك ما يأتي:

- في الدراسات الطبية،فان مشاهدات ضغط الدم تمثل المتغير X ومعدل النبض يمثل المتغير Y لكل مجموعة من المرضى.في هذه الحالة يكون كلا المتغيرين هما متغيرات كمية مستمرة.
- في شركة ما ، فان X يمثل العدد الشهري للسفرات وان Y يمثل مصاريف الانتقال.في مثل هذه الحالة فان X يمثل متغير كمي متقطع بينما Y يمثل متغير كمي مستمر.

في هذا الفصل سوف نوسع مفهوم المتغير العشوائي ودالة التوزيع الاحتمالي له إلى مفهوم المتغيرين العشوائيين ودوالهما التوزيعية الاحتمالية المشتركة(Joint Probability Distribution) ، والفكرة المهمة التي سوف نقدمها هي فكرة الاستقلال(Independence) ،الارتباط(Correlation) والتوقع الشرطي (Conditional Expectation) وسوف ندرس بعض التوزيعات الثنائية المتقطعة والمستمرة .

2-7 المتغير العشوائي المتقطع الثنائي ودوال التوزيع الاحتمالي

Bivariate Discrete Random Variables And Joint Probability Distribution Functions

لكي نتناول المتغير العشوائي المتقطع الثنائي ودالة التوزيع الاحتمالي المشتركة المرتبطة به،سوف نتناول بعض الأمثلة البسيطة.

مثال 7-1

حاوية تحتوي على خمسة مواد،اثنان منها سليمة بشكل كامل ويرمز لها بالرمز S_1 ، S_2 ،واثنان منها تحتوي على عيب بسيط ويرمز لها بالرمز M_1 ، M_2 ، والباقي فيها عيب رئيسي ويرمز لها بالرمز F .سحبت منها عينة عشوائية مكونة من مادتين وبدون إرجاع.ليكن كل من X و Y متغيرين عشوائيين يمثلان عدد المواد السليمة والمواد التي فيها عيب بسيط على التوالي في العينة .

في بداية الأمر نقوم بحساب عناصر فضاء العينة S وكما يأتي:

$$S = \{S_1S_2, S_1M_1, S_1M_2, S_1F, S_2M_1, S_2M_2, S_2F, M_1M_2, M_1F,$$
$$M_2F, S_2S_1, M_1S_1, M_2S_1, FS_1, M_1S_2, M_2S_2, FS_2, M_2M_1,$$
$$FM_1, FM_2\}$$

وبهذا يكون عدد العناصر لفضاء العينة هو (20) عنصراً يمكن حسابها باستخدام علاقة التباديل.ويمكن أن نضع القيم المرتبطة لـ(X,Y) بالأحداث اعلاه ونمثلها بـ(x,y) وكما يأتي:

$$(x,y)=(2,0)\ (1,1)\ (1,1)\ (1,0)\ (1,1)\ (1,1)\ (1,0)\ (0,2)\ (0,1)\ (0,1)$$
$$(2,0)\ (1,1)\ (1,1)\ (1,0)\ (1,1)\ (1,1)\ (1,0)\ (0,2)\ (0,1)\ (0,1)$$

ولاننا أستخدمنا عينة عشوائية بدون إرجاع(without replacement) فأن كل حدث بسيط يكون متساوي الحدوث(equally likely to occur) ويساوي $\frac{1}{20}$ ولهذا سوف نحصل على التوزيع الاحتمالي المشترك(joint probability distribution) لـ X و Y:

$$
\begin{array}{llllll}
(x\,,y) & :(2,0) & (1,1) & (1,0) & (0,1) & (0,2) \\
P(X=x,\ Y=y): & 2/20 & 8/20 & 4/20 & 4/20 & 2/20
\end{array}
$$

وعلى أساس هذا المثال سوف نضع بعض التعاريف الآتية:

تعريف 7-1

ليكن S فضاء عينة لتجربة عشوائية ما،وليكن كل من $X=X(e)$ و $Y=Y(e)$ دالتي القيمة الحقيقية التي تربط عدد حقيقي إلى كل عنصر e من عناصر فضاء العينة S .لذا فان الزوج المرتب (X,Y) يسمى متغير عشوائي ذات البعدين(وأحيانا يسمى متجه عشوائي).أذا كان عدد القيم الممكنة لـ (X,Y) منتهي(infinite) أو معدود غير منتهي(countably infinite) فان (X,Y) يسمى متغير عشوائي متقطع ذات البعدين.

سوف نمثل القيم الممكنة لـ (X,Y) بـ (x_i,y_j) حيث أن $i=1,2,...,m_x$ و $j=1,2,...,m_y$ حيث أن m_x و m_y يذهبان إلى ∞ .

وان $R_{x,y}=\{(x_i,y_j):i=1,2,...,m_x\,;j=1,2,...,m_y\}$ يمثل فضاء ذات البعدين لـ (X,Y) .كما حصل في حالة البعد الواحد. سوف نهتم بالاحتمالات مثل $P(X=x,Y=y)$ و $P(a_1\le X\le a_2,b_1\le Y\le b_2)$.إن الدالة الرياضية التي نستخدمها لحساب هذه الاحتمالات تسمى دالة التوزيع الاحتمالي المشتركة (joint probability distribution function) ويرمزلها بالرمز $(j.p.d.f)$ للمتغيرين العشوائيين X و Y .

تعريف 7-2

ليكن (X,Y) متغير عشوائي متقطع ذات بعدين معرف على الفضاء $R_{x,y}$.

فان $P(X=x,Y=y)$ يرمز لها بالرمز $f(x,y)$ وتسمى دالة التوزيع الاحتمالي المشتركة لـ X و Y وتحقق الخواص الآتية:

1. $0 \le f(x,y) \le 1$.

2. $\sum\limits_{(x,y)\in R_{x,y}} \sum f(x,y) = 1$

3. $P\{(X,Y)\in A\} = \sum\limits_{(x,y)\in R_{x,y}} \sum f(x,y)$

حيث A هي مجموعة جزئية من الفضاء $R_{x,y}$.

مثال 7-2

في المثال 7-1 نفرض أن العينة المختارة تتكون من ثلاث مواد تسحب بشكل عشوائي من الحاوية ،فإذا كان كل من X و Y متغيرين عشوائيين متقطعين يمثلان المواد السليمة والمواد التي تحوي على عطل بسيط على التوالي،إحسب دالة التوزيع الاحتمالي المشتركة لـ X و Y ؟

الحل

بإهمال عملية الترتيب فان الأحداث الممكنة للعينة هي:

$S_1S_2M_1 \quad S_1S_2M_2 \quad S_1S_2F \quad S_1M_1M_2 \quad S_1M_1F \quad S_1M_2F$

$S_2M_1M_2 \quad S_2M_1F \quad S_2M_2F \quad M_1M_2F$

وان قيم (x,y) بالنسبة لـ (X,Y) المقابلة لهاهي:

$(x,y)=(2,1)\ (2,1)\ (2,0)\ (1,2)\ (1,1)\ (1,1)\ (1,2)\ (1,1)\ (1,1)\ (0,2)$

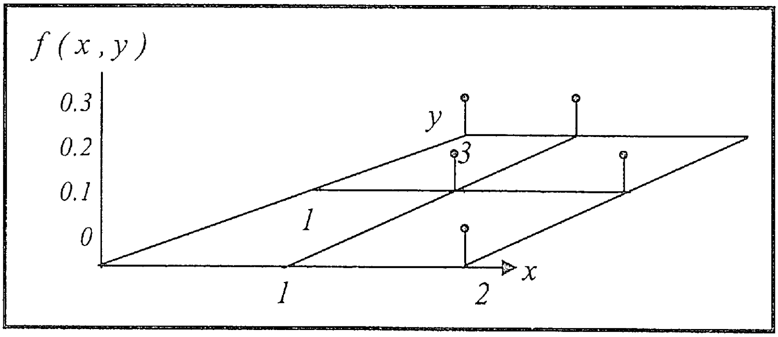
وان الاختيار العشوائي للنتائج الممكنة اعلاه يكون باحتماليات متساوية أي أن الاحتمالية $1/10$ ،لذلك فان دالة التوزيع الاحتمالية المشتركة (j.p.d.f) للمتغيرين العشوائيين X و Y هي:

(x,y)	$(2,1)$	$(2,0)$	$(1,1)$	$(0,2)$	$(1,2)$
$P(X=x,\ Y=y)$	$2/10$	$1/10$	$4/10$	$1/10$	$2/10$

وعلية يمكن حساب الاحتمال الآتي:

$$P(X \leq 1, Y \geq 1) = P(X = 0, Y = 2) + P(X 1, Y = 1) + P(X = 1, Y = 2)$$
$$= 7/10$$

ويمكن تمثيل الدالة بالرسم المبين في الشكل7-1 أدناه:



الشكل 7-1 (رسم دالة الاحتمال التوزيعية المشتركة لـ X و Y)

لقد لاحظنا في الفصول السابقة بان دالة التوزيع التجميعية (c.d.f)تلعب دوراً مهماً في دراسة المتغير العشوائي ذات البعد الواحد،ولذلك سوف نوسع هذا المفهوم بالنسبة للمتغيرات العشوائية المتقطعة ذات البعدين من خلال التعريف الآتي:

تعريف 7-3

ليكن (X,Y) متغيرين عشوائيين متقطعين ، فان دالة التوزيع التجميعية الثنائية لـ X و Y (bivariate c.d.f)تعرف كما يأتي:

$$F(x,y) = P(X \leq x, Y \leq y) = \sum_{u \leq x} \sum_{v \leq y} f(u,v) \quad \ldots\ldots\ldots\ldots\ldots\ldots(7\text{-}1)$$

ودالة التوزيع التجميعية الثنائية تحقق عدداً من الخواص ،علماً أن الخواص الثلاثة الأولى منها تبرهن بأسلوب مشابه لما تم برهانه في حالة المتغير الواحد.

والنظرية الآتية تُبين هذه الخواص حيث سنبرهن منها الفقرة (d)فقط.

نظرية 7-1

لتكن $F(x,y)$ دالة التوزيع التجميعية الثنائية لـ(X,Y) فان $F(x,y)$ تحقق الخواص التالية:

a) $F(\infty,\infty)=1$

b) $F(-\infty,y)=F(x,-\infty)=0$

c) $F(x,y)$ دالة غير متناقصة في كل متغير متقطع.

d) لجميع القيم $x_1,x_2(x_2>x_1)$ و $y_1,y_2(y_2>y_1)$، يكون:

$$F(x_2,y_2)-F(x_1,y_2)-F(x_2,y_1)+F(x_1,y_1)\geq 0$$

البرهان

من الشكل 7-2 أدناه نلاحظ أن:

$$\{(X\leq x_2)\cap(Y\leq y_2)=\{(X\leq x_1)\cap(Y\leq y_1)\}\cup\{(X\leq x_1)\cap(y_1<Y\leq y_2)\}$$
$$\cup\{(x_1<X\leq x_2)\cap(Y\leq y_1)\}\cup\{(x_1<X\leq x_2)\cap(y_1<Y\leq y_2)\}$$

لذلك سوف نستخدم قانون الجمع للأحداث المنفصلة فنحصل على:

$$P(x_1<X\leq x_2,y_1<Y\leq y_2)=P(X\leq x_2,Y\leq y_2)$$
$$-P(X\leq x_1,Y\leq y_1)\ \ -P(X\leq x_1,y_1<Y\leq y_2)$$
$$-P(x_1<X\leq x_2,Y\leq y_1)\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7\text{-}2)$$

نكتب:

$$P(X\leq x_1,y_1<Y\leq y_2)=P(X\leq x_1,Y\leq y_2)-P(X\leq x_1,Y\leq y_1)$$

و

$$P(x<X\leq x_2,Y\leq y_1)=P(X\leq x_2,Y\leq y_1)-P(X\leq x_1,Y\leq y_1)$$

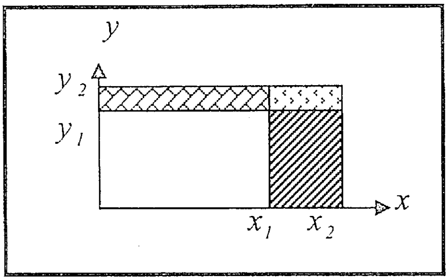
نعوض عن العلاقتين اعلاه في العلاقة $(7-2)$ فنحصل على :

$$P(x_1<X\leq x_2,y_1<Y\leq y_2)=P(X\leq x_2,Y\leq y_2)$$
$$-P(X\leq x_1,Y\leq y_2)\ -P(X\leq x_2,Y\leq y_1)+P(X\leq x_1,Y\leq y_1)$$

$$= F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \quad \ldots\ldots\ldots\ldots (7\text{-}3)$$

وبهذا نحصل على:

$$P(x_1 < X \leq x_2, y_1 < Y \leq y_2) \geq 0$$



الشكل (7-2)

مثال 7-3

اذا كانت دالة التوزيع الاحتمالي المشتركة(j.p.d.f) لـ X و Y هي:

$$f(x, y) = \begin{cases} \binom{2}{y}\left(\frac{1}{2}\right)^{x+2} & , x = 1, 2, \ldots, \infty \quad ; y = 0, 1, 2. \\ 0 & otherwise \end{cases}$$

حيث $1 \leq x < \infty$ و $0 \leq y \leq 2$، أوجد دالة التوزيع التجميعية الثنائية لـ (X,Y)؟

الحل

نحسب دالة التوزيع التجميعية كما يأتي:

$$F(x, y) = \sum_{u=1}^{x} \sum_{v=0}^{y} f(u, v) = \sum_{u=1}^{x} \left(\frac{1}{2}\right)^{u+2} \sum_{v=0}^{y} \binom{2}{v}$$

$$= \frac{1}{4}\left\{1 - \left(\frac{1}{2}\right)^{x}\right\} \sum_{v=0}^{y} \binom{2}{v}$$

وعليه تكون الدالة التجميعية هي:

$$F(x,y)=\begin{cases} 0 & for \quad x<1 \quad and \quad y<0 \\ \frac{1}{4}\{1-(\frac{1}{2})^x\}\sum_{v=0}^{y}\binom{2}{v} & for \quad x\geq 1, 0\leq y\leq 2 \\ 1-(\frac{1}{2})^x & for \quad x\geq 1, y>2 \end{cases}$$

إن الدالة اعلاة تحقق الخواص في نظرية(7-1) وعلى الطالب إثبات ذلك.

7-3 المتغير العشوائي المستمر الثنائي ودالة الكثافة الاحتمالية المشتركة

Bivariate Continuous Random Variables And Joint Probability Density Functions:

لإثارة فكرة التوزيع الثنائي المستمر،فإننا سنوضح مفهوم المدرج التكراري النسبي (relative frequency histogram) للتعامل مع هذه الحالة.

نفرض أن لدينا n من أزواج القيم $(x_i, y_i), i=1,2,...,n$، فإذا كانت n كبيرة نسبياً ،فإننا نضع المشاهدات في خلايا أو فترات صفية ثنائية هي $(c_0,c_1),(c_1,c_2),...,(c_{k-1},c_k)$ بالنسبة لمحور x والفترات الصفية $(d_0,d_1),(d_1,d_2),...,(d_{l-1},d_l)$ بالنسبة لمحور y .فإذا كان f_{ij} يمثل التكرار أو عدد المشاهدات في الخلية (i,j) المعرفة بالفترة الصفية الثنائية (c_{i-1},c_i) و (d_{j-1},d_j) فان f_{ij}/n يمثل التكرار النسبي للمشاهدات في الخلية (i,j) حيث أن $i=1,2,...,k$ و $j=1,2,...,l$.

ويرسم المدرج التكراري النسبي من خلال رسم الأعمدة فوق الخلايا بحيث أن حجم العمود فوق الخلية (i,j) يساوي التكرار النسبي f_{ij}/n .وهكذا فان المدرج التكراري النسبي يمثل الدالة $h_n(x,y)$ المعرفة بالصيغة الآتية:

$$h_n(x,y)=\frac{f_{ij}}{n(c_i-c_{i-1})(d_j-d_{j-1})} \qquad ....(7-4)$$

بالنسبة لـ $c_{i-1}<x\leq c_i, d_{j-1}<y\leq d_j \quad, i=1,2,...,k, j=1,2,...,l$

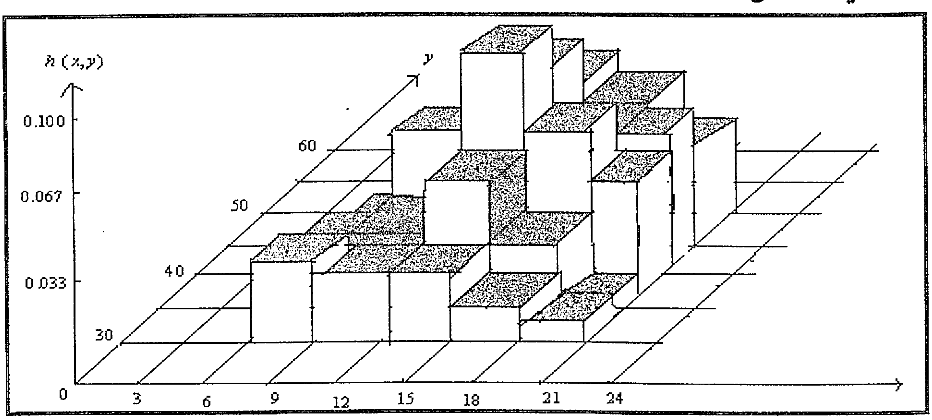
مثال 7-4

البيانات في الجدول (7-1) أدناه تبين لنا ترتيب التوزيع التكراري الثنائي لمختلف سنوات الخدمة x و الراتب y في وحدات مرمزة لعينة من (150)عامل الذين يعملون في شركة ما، حيث أكملوا خدمة لاتقل عن ست سنوات في الشركة.

الجدول (7-1)

الراتب y \ x	من 30 وتحت 35	من 35 وتحت 40	من 40 وتحت 45	من 45 وتحت 50	من 50 وتحت 55
6 وأقل من 9	*5*	*4*	*3*	*3*	*1*
9 وأقل من 12	*4*	*5*	*10*	*5*	*6*
12 وأقل من 15	*2*	*8*	*15*	*13*	*10*
15 وأقل من 18	*2*	*4*	*10*	*8*	*8*
18 وأقل من 21	*1*	*1*	*7*	*8*	*5*

حيث x يمثل سنوات الخدمة.ويمكن رسم المدرج التكراري النسبي للجدول أعلاه كما في الشكل 7-3 :



الشكل (7-3) يبين المدرج التكراري النسبي الثنائي

نلاحظ أن دالة المدرج التكراري النسبي الثنائي $h_n(x,y)$ تحقق الخواص الآتية:

a) $h_n(x,y) \geq 0$ لجميع x و y.

b) الحجم الكلي المحدد بالمستوي (x,y) من الأعلى والدالة $h_n(x,y)$ من الأسفل يساوي واحد أي أن:

$$\int_{c_0}^{c_k}\int_{d_0}^{d_l} h_n(x,y)\,dxdy = 1$$

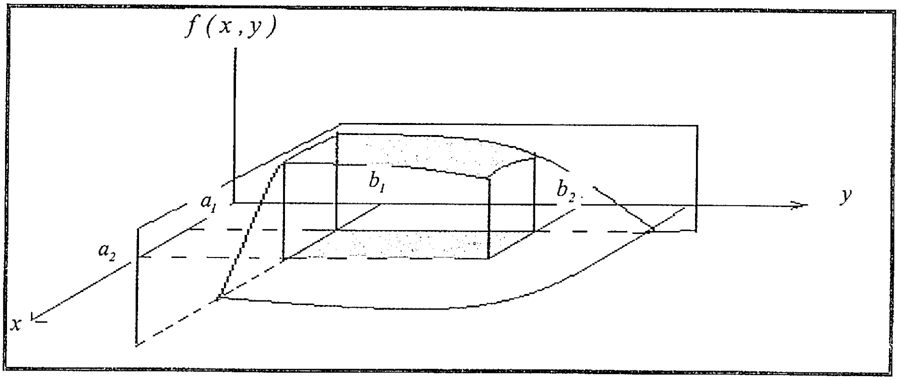
c) احتمالية أي حدث A مكون من اتحاد الفترات الصفية الثنائية يمكن أن نحسبه من العلاقة الآتية:

$$p(A) \approx \iint_A h_n(x,y)\,dxdy$$

ولتوضيح ذلك أكثر،نفرض أن عدد المشاهدات n يزداد وبذلك نستطيع خفض عدد الخلايا وبأخذ الغاية للدالة $h_n(x,y)$ فإنها تقترب تدريجياً من الدالة الرياضية التي نسميها $f(x,y)$ والتي تعطي القيمة الحقيقية للاحتمال المتعلق بالمتغيرين X و Y من خلال التكامل الآتي:

$$P(a_1 < X < a_2, b_1 < Y < b_2) = \int_{a_1}^{a_2}\int_{b_1}^{b_2} f(x,y)\,dydx \quad \ldots\ldots\ldots\ldots\ldots(7\text{-}5)$$

إن هذه الاحتمالية تمثل حجم الجسم فوق المستطيل المظلل في الشكل (7-4) أدناه:



الشكل(7-4)

وعلى ضوء هذه المفاهيم سنضع التعريف الآتي:

تعريف7-4

ليكن (X,Y) متغيراً عشوائياً ذو بعدين الذي نفترض إن جميع قيمه في مجموعة جزئية غير معدودة من الفضاء الاقليدي، فان مثل هذا المتغير يسمى متغير عشوائي مستمر ذات البعدين ،ومجموعة القيم التي يأخذها المتغير (X,Y) تسمى مدى الفضاء ويرمز لها بالرمز $R_{X,Y}$.[4]

إن الدالة $f(x,y)$ تسمى دالة الكثافة الاحتمالية المشتركة (joint probability density function) للمتغيرين X و Y ويرمز لها بالرمز(j.p.d.f) اذا حققت الشروط الآتية:

1) $f(x,y) \geq 0$ لجميع قيم $(x,y) \in R_{X,Y}$.

2) $\iint_{R_{X,Y}} f(x,y)\,dx\,dy = 1$

3) احتمالية أي حدث $(X,Y) \in A$ ،عندما A مجموعة جزئية من $R_{X,Y}$ هي:

$$P\{(X,Y) \in A\} = \iint_A f(x.y)\,dx\,dy$$

ملاحظة:

لأي عددين موجبين Δx و Δy صغيران جداً،فان $f(x,y)\Delta x\Delta y$ تساوي تقريباً $P(x \leq X \leq x+\Delta x, y \leq Y \leq y+\Delta y)$.

مثال7-5

لتكن الدالة

$$f(x,y) = \begin{cases} \frac{1}{8}(6-x-y), & 0 < x < 2, 2 < y < 4 \\ 0, & otherwise \end{cases}$$

بين إنها دالة كثافة احتمالية مشتركة؟ثم احسب $p(X < 1, Y > 3)$ ؟

الحل

للتحقق من كون الدالة هي دالة كثافة احتمالية يجب أن نحقق الشروط في التعريف 7-4 حيث انه من الواضح بان الدالة $f(x,y)\geq 0$ لجميع قيم x و y في الفترة المعرفة عليها الدالة،كذلك فان :

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_{0}^{2}\int_{2}^{4}\frac{1}{8}(6-x-y)\,dy\,dx$$

$$= \int_{0}^{2}\left[\frac{6y}{8}-\frac{xy}{8}-\frac{y^2}{16}\right]_{2}^{4} dx$$

$$= \int_{0}^{2}\left[\frac{3}{4}-\frac{x}{4}\right] dx = 1$$

وهذا يبين أن الدالة $f(x,y)$ هي دالة الكثافة الاحتمالية.ولحساب

$$P(X<1,Y>3) = \int_{0}^{1}\int_{3}^{4}\frac{1}{8}(6-x-y)\,dy\,dx$$

$$= \int_{0}^{1}\left[\frac{3y}{4}-\frac{xy}{8}-\frac{y^2}{16}\right]_{3}^{4} dx$$

$$= \int_{0}^{1}\left[\frac{5}{16}-\frac{x}{8}\right] dx = \frac{1}{4}$$

7-3-1 دالة التوزيع التجميعية المشتركة

The Joint Commutative Distribution Function

للحصول على دالة التوزيع التجميعية المشتركة *(j.c.d.f)* في حالة المتغير العشوائي المستمر الثنائي (X,Y) فإننا فقط نستبدل عملية الجمع المضاعف في حالة المتغير المتقطع إلى عملية التكامل المضاعف.

وعلية اذا كان (X,Y) متغير عشوائي مستمر ذات بعدين فنعرف دالة التوزيع التجميعية المشتركة كالآتي:

$$F(x,y) = P(X\leq x, Y\leq y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,dv\,du \;\ldots\ldots\ldots\ldots(7\text{-}6)$$

ملاحظة:

إن الدالة $F(x,y)$ تحقق جميع الخواص التي ورد ذكرها في نظرية (7-1)عندما تكلمنا عن حالة المتغير العشوائي المتقطع الثنائي.وبالإضافة إلى ذلك فانه اذا كانت المشتقة الأولى والثانية للدالة $F(x,y)$ موجودة فان

$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y)$ أي أننا نحصل منها على دالة الكثافة الاحتمالية المشتركة للتوزيع.ويمكن توضيح ذلك من خلال:

$$\frac{\partial^2 F(x,y)}{\partial x\,\partial y} = \frac{\partial}{\partial x}\left\{ \lim_{\Delta y \to 0} \frac{F(x,y+\Delta y)-F(x,y)}{\Delta x\,\Delta y} \right\}$$

$$= \lim_{\Delta x \to 0} \lim_{\Delta y \to 0} \frac{\left[F(x+\Delta x, y+\Delta y)-F(x+\Delta x, y)\right]-\left[F(x,y+\Delta y)-F(x,y)\right]}{\Delta x\,\Delta y}$$

ومن العلاقة(7-3) نحصل على: $\lim_{\Delta x \to 0}$

$$\lim_{\Delta x \to 0} \lim_{\Delta y \to 0} \frac{\left[p(x < X \le x+\Delta x,\ y < Y \le y+\Delta y)\right]}{\Delta x\,\Delta y} = f(x,y)$$

مثال 7-6

لتكن الدالة في المثال (7-5):

$$f(x,y) = \begin{cases} \frac{1}{8}(6-x-y), & 0 < x < 2,\ 2 < y < 4 \\ \\ 0 \quad , & otherwise \end{cases}$$

أحسب دالة التوزيع التجميعية المشتركة$(j.c.d.f)$؟

الحل

حسب التعريف(7-4)فإن: $F(x,y)=P(X\le x,Y\le y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dvdu$

بالنسبة إلى $x\le 0$ أو $y\le 2$ و $f(u,v)=0$ بالنسبة إلى $u\le x,t\le y$ وكذلك فإن $F(x,y)=0$ بالنسبة إلى $x\le 0,y\le 2$.

وبالنسبة إلى $0<x<2,2<y<4$ نحصل على: $F(x,y)=\int_{0}^{x}\int_{2}^{y}\frac{1}{8}(6-u-v)dvdu$

$$F(x,y)=\int_{0}^{x}\frac{1}{8}\left\{6-u)(y-2)-\frac{y^2}{2}+2\right\}du$$

$$F(x,y)=\frac{1}{16}x(y-2)(10-y-x)$$

كذلك نلاحظ أن :

$$F(x,y)=\int_{0}^{x}\int_{2}^{4}\frac{1}{8}(6-u-v)dvdu=\frac{1}{8}x(6-x)$$

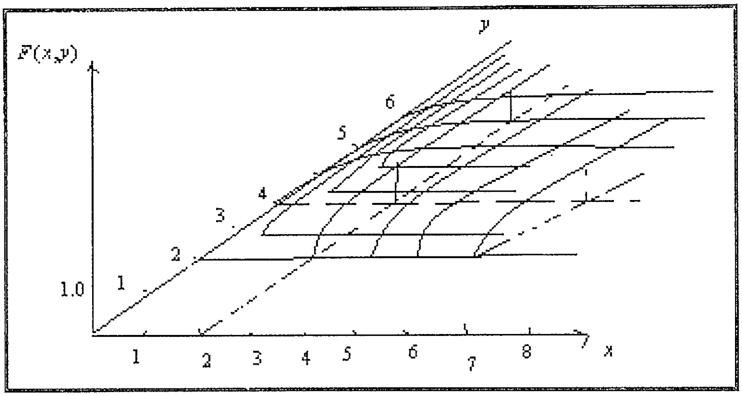
اذا كانت $0<x<2,y\ge 4$.وبأسلوب مشابه فان :

$$F(x,y)=\int_{0}^{2}\int_{2}^{y}\frac{1}{8}(6-u-v)dvdu=\frac{1}{8}(y-2)(8-y)$$

إذا كانت $x\ge 2,2<y<4$. وأخيرا نحصل على:

$F(x,y)=\int_{0}^{2}\int_{2}^{4}\frac{1}{8}(6-u-v)dvdu=1$ حيث أن $x\ge 2,y\ge 4$.

والشكل (7-5)أدناه يبين رسم للدالة أعلاه:



الشكل (7-5)

7-4 التوزيعات الهامشية والشرطية

Marginal And Conditional Distributions

لتكن $f(x,y)$ هي دالة التوزيع الاحتمالية المشتركة للمتغيرين العشوائيين X و Y. فأحيانا نهتم بدالة التوزيع الاحتمالي $f_1(x)$ للمتغير العشوائي X، أو دالة التوزيع الاحتمالي $f_2(y)$ للمتغير العشوائي Y. إن الدالتين المذكورتين تشيران إلى الدالتين الهامشيتين للمتغيرين العشوائيين X و Y على التوالي.

ومن تعريف دالة التوزيع التجميعية نستطيع أن نوجد دالة التوزيع التجميعية الهامشية للمتغيرين X و Y.

مثال7-7

لتكن دالة التوزيع الاحتمالية المشتركة للمتغيرين العشوائيين المتقطعين X و Y هي:

x/y	1	2	3	Total
0	0.3	0.2	0.2	0.7
1	0.0	0.1	0.1	0.2
2	0.1	0.0	0.0	0.1
Total	0.4	0.3	0.3	1.0

أحسب دالة التوزيع الاحتمالية الهامشية المشتركة ودالة التوزيع التجميعية للمتغيرين العشوائيين X و Y؟

الحل

لحساب الدالة الهامشية:

$$f_1(0) = P(X = x) = P(X = 0, Y = 1) + P(X = 0, Y = 2)$$
$$P(X = 0, Y = 3) = 0.3 + 0.2 + 0.2 = 0.7$$

$$f_1(1) = P(X = 1, Y = 1) + P(X = 1, Y = 2) + P(X = 1, Y = 3)$$
$$= 0.0 + 0.1 + 0.1 = 0.2$$

$$f_1(1) = P(X = 2, Y = 1) + P(X = 2, Y = 2) + P(X = 2, Y = 3)$$
$$= 0.1 + 0.0 + 0.0 = 0.1$$

وبنفس الأسلوب فان:

$$f_2(1) = P(Y = y) = P(X = 0, Y = 1) + P(X = 1, Y = 1)$$
$$P(X = 2, Y = 1) = 0.3 + 0.0 + 0.1 = 0.4$$

$$f_2(2) = P(X = 0, Y = 2) + P(X = 1, Y = 2) + P(X = 2, Y = 2)$$
$$= 0.2 + 0.1 + 0.0 = 0.3$$

$$f_2(3) = P(X = 0, Y = 3) + P(X = 1, Y = 3) + P(X = 2, Y = 3)$$
$$= 0.2 + 0.1 + 0.0 = 0.3$$

ولحساب دالة التوزيع التجميعية للدالتين $f_1(x)$ و $f_2(y)$ للمتغيرين X و Y وكما يلي:

x :	0	1	2
$f_1(x)$:	0.7	02	0.1
$F_1(x)$:	0.7	0.9	1.0
y :	1	2	3
$f_2(y)$:	0.4	0.3	0.3
$F_2(y)$:	0.4	0.7	1.0

من المثال السابق يمكن أن نضع التعريف الآتي:

تعريف7-5

ليكن X و Y متغيرين عشوائيين متقطعين لهما دالة التوزيع الاحتمالية المشتركة $f(x,y)$ في الفضاء $R_{X,Y}$. فان دالتي التوزيع الاحتمالية الهامشية للمتغيرين X و Y هما:

$$f_2(y) = \sum_{x(y)} f(x,y) \text{ و } f_1(x) = \sum_{y(x)} f(x,y) \quad \ldots\ldots\ldots\ldots\ldots\ldots(7\text{-}7)$$

حيث أن $x, y \in R_{X,Y}$.

عندما $\sum_{x(y)}$ يمثل مجموع جميع القيم الممكنة بالنسبة إلى x المرتبطة مع قيم y المعطاة،وان $\sum_{y(x)}$ يمثل مجموع جميع القيم الممكنة بالنسبة إلى y المرتبطة مع قيم x المعطاة.

وأن الدالتين التجميعيتين الهامشيتين بالنسبة إلى X و Y تعطى بالشكل الآتي:

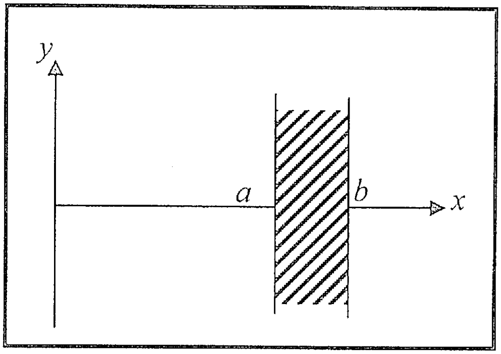
$$F_2(y) = f(\infty, y) = \sum_{v \le y} f_2(v) \text{ و } F_1(x) = f(x,\infty) = \sum_{u \le x} f_1(u) \quad \ldots.(7\text{-}8)$$

على التوالي.

ملاحظة:

يبدو للقارئ بأن كل من الدالتين $f_1(x)$ و $f_2(y)$ ليستا دوال ثنائية ،ولكن سميت دوال توزيع احتمالي هامشي، لأنها أشتقت من الدالة الثنائية $f(x,y)$. [6]

سوف نعرّف التوزيعات الهامشية في حالة المتغيرات المستمرة بنفس الأسلوب الذي عرفنا به التوزيعات الهامشية في حالة المتغيرات المتقطعة ولكن من خلال استبدال المجموع بعملية التكامل.ولهذا سنفرض $f(x,y)$ بأنها دالة التوزيع الاحتمالية المشتركة للمتغيرين العشوائيين المستمرين X و Y،وسوف نهتم بالاحتمال $P(a < X < b)$ حيث أنه يمثل الحجم تحت الدالة $f(x,y)$ والشريط المظلل في الشكل (7-6) أدناه:



الشكل(7-6)

نلاحظ أن:

$$P(a<X<b)=\int_{a}^{b}\int_{-\infty}^{\infty}f(x,y)\,dy\,dx=\int_{a}^{b}f_1(x)\,dx$$

عندما $f_1(x)=\int_{-\infty}^{\infty}f(x,y)\,dy$.

وبنفس الأسلوب بالنسبة للدالة الهامشية $f_2(y)$.

تعريف 7-6

ليكن X و Y متغيرين عشوائيين مستمرين لهما دالة التوزيع الاحتمالية المشتركة $f(x,y)$ في الفضاء $R_{X,Y}$. فان دالتي التوزيع الاحتمالية الهامشية للمتغيرين X و Y هما:

$$f_2(y)=\int_{-\infty}^{\infty}f(x,y)\,dx \text{ و } f_1(x)=\int_{-\infty}^{\infty}f(x,y)\,dy \quad\ldots\ldots\ldots\ldots\ldots(7\text{-}9)$$

حيث أن $x,y\in R_{X,Y}$.

وأن الدالتين التجميعيتين الهامشيتين بالنسبة إلى X و Y تعطى بالشكل الآتي:

$$F_2(y)=F(\infty,y)=\int_{-\infty}^{y}f_2(v)\,dv \text{ و } F_1(x)=F(x,\infty)=\int_{-\infty}^{x}f_1(u)\,du \quad\ldots\ldots(7\text{-}10)$$

على التوالي.

مثال 7-8

لتكن دالة التوزيع الاحتمالية المشتركة للمتغيرين العشوائيين X و Y هي:

$$f(x,y)=\begin{cases} x^2y^2/81 & ,0<x<3,0<y<3 \\ 0 & Elsewhere \end{cases}$$

أوجد دالة التوزيع الاحتمالية الهامشية بالنسبة إلى X و Y ؟ ثم أوجد دواليهما التجميعية ،ودالة التوزيع التجميعية المشتركة ؟

الحل

دوال التوزيع الاحتمالية الهامشية هي:

$$f_1(x)=\int_{-\infty}^{\infty} f(x,y)dy=\int_0^3 \frac{x^2y^2}{81}dy=\frac{x^2y^3}{81\times 3}\Big|_0^3=\frac{x^2}{9}$$

$$f_2(y)=\int_{-\infty}^{\infty} f(x,y)dx=\int_0^3 \frac{x^2y^2}{81}dx=\frac{x^3y^2}{81\times 3}\Big|_0^3=\frac{y^2}{9}$$

دوال التوزيع التجميعية الهامشية هي :

$$F_1(x)=\int_{-\infty}^{x} f_1(u)du=\int_0^x \frac{u^2}{9}du=\frac{u^3}{9\times 3}\Big|_0^x=\frac{x^3}{27} \quad 0<x<3$$

$$F_2(y)=\int_{-\infty}^{x} f_2(v)dv=\int_0^x \frac{v^2}{9}dv=\frac{v^3}{9\times 3}\Big|_0^y=\frac{y^3}{27} \quad 0<y<3$$

نلاحظ أن :

$F_1(x)=\begin{cases} 1 & ,x>3 \\ 0 & ,x\le 0 \end{cases}$ وكذلك $F_2(y)=\begin{cases} 1 & ,y>3 \\ 0 & ,y\le 0 \end{cases}$ وبذلك تكون دالة التوزيع التجميعية المشتركة للمتغيرين X و Y هي:

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\,dv\,du=\begin{cases}0 & ,x<0 \quad or \quad y<3\\ x^3y^3/729 & ,0<x<3,0<y<3\\ x^3/27 & ,0<x<3,y>3\\ y^3/27 & ,x>3,0<y<3\\ 1 & ,x>3,y>3\end{cases}$$

1-4-7 التوزيع الشرطي الثنائي Bivariate Conditional Distribution

يستخدم هذا المفهوم بشكل كبير في الإحصاء الرياضي ،ولقد سبق لنا وأن ناقشنا في الفصل الثاني موضوع الاحتمال الشرطي لأي حدث A عندما يكون الحدث B معطى أو حاصل فان :

$P(A/B)=\dfrac{P(A\cap B)}{P(B)}$ و $P(B/A)=\dfrac{P(A\cap B)}{P(A)}$ حيث أن كل من $P(A)\neq 0$ و $P(B)\neq 0$.

فإذا كان المتغيران العشوائيان المتقطعان X و Y يقابلان الحدثين A و B على التوالي فعندها يكون بالإمكان التحدث عن دالة التوزيع الاحتمالي الشرطية للمتغير العشوائية المتقطعة.[5]

2-4-7 دالة التوزيع الاحتمالي الشرطية للمتغيرات العشوائية المتقطعة

Conditional Probability Distribution Function Of Discrete Random Variables:

تعريف 7-7

ليكن X و Y متغيرين عشوائيين متقطعين لهما دالة التوزيع الاحتمالي المشتركة (joint $p.d.f$) $f(x,y)$ ودالتي التوزيع الاحتمالي الهامشية (marginal $p.d.f$) $f_1(x)$ و $f_2(y)$ على التوالي حيث كلاهما تمثلان دوال موجبة فان :

$$P(X=x/Y=y)=\frac{P(X=x,Y=y)}{P(Y=y)}=\frac{f(x,y)}{f_2(y)}$$

نضع $g_1(x/y)=P(X=x/Y=y)$ بحيث أن:

$$g_1(x/y)=\frac{f(x,y)}{f_2(y)} \quad \text{.......................(7-11)}$$

حيث $f_2(y)>0$.

فإن $g_1(x/y)$ تسمى دالة التوزيع الاحتمالي الشرطية للمتغير العشوائي المتقطع X عندما يكون المتغير العشوائي المتقطع $Y=y$ معطى.

وبنفس الأسلوب نحصل على دالة التوزيع الاحتمالي الشرطية للمتغير العشوائي المتقطع Y عندما $X=x$ معطى وهي:

$$g_2(y/x)=\frac{f(x,y)}{f_1(x)} \quad \text{.......................(7-12)}$$

حيث $f_1(x)>0$.

نلاحظ أنّ كل من الدالتين $g_1(x/y)$ و $g_2(y/x)$ تحققان شروط دوال التوزيع الاحتمالي المتقطع وهي:

- $g_1(x/y)>0$ و $g_2(y/x)>0$.
- كذلك فان : $\sum_{\forall x} g_1(x/y)=\frac{\sum_{\forall x} f(x,y)}{f_2(y)}=\frac{f_2(y)}{f_2(y)}=1$

$$\sum_{\forall y} g_2(y/x)=\frac{\sum_{\forall y} f(x,y)}{f_1(x)}=\frac{f_1(x)}{f_1(x)}=1$$

مثال 7-9

اذا كان الطلب الأسبوعي على المنتجين A و B ، الذي يباع في الأسواق ،لهما دالة التوزيع الاحتمالي المشتركة ، قيمها كما مبينة في الجدول أدناه.أحسب دالة التوزيع الاحتمالي الشرطية لطلب المنتج A عندما يكون الطلب على المنتج B معطى وهو (1) وحدة،كذلك أحسب دالة التوزيع الاحتمالية التجميعية لطلب المنتج B عندما يكون الطلب على A هو (2) وحدة.

		A				
	x \ y	0	1	2	3	total
B	0	.05	.10	.15	.05	.35
	1	.10	.20	.10	.05	.45
	2	.05	.10	.05	.00	.20
	total	.20	.40	.30	.10	1.0

الحل

ليكن X متغيراً عشوائياً يمثل الطلب على المنتج A و Y متغير عشوائي يمثل الطلب على المنتج B.

ولتكن $g_1(x/y)$ تمثل دالة التوزيع الشرطي إلى X عندما $Y=y$ معطى و $g_2(y/x)$ تمثل دالة التوزيع الشرطي إلى Y عندما $X=x$ معطى.

نستخدم العلاقتين $(5-11)$ و $(5-12)$ لحساب دالة التوزيع الاحتمالي الشرطية إلى X عندما يكون $Y=1$ وهي:

$$g_1(0/1)=f(0,1)/f_2(1)=(0.10)/(0.45)=0.22$$

$$g_1(1/1)=f(1,1)/f_2(1)=(0.20)/(0.45)=0.44$$

$$g_1(2/1)=f(2,1)/f_2(1)=(0.10)/(0.45)=0.22$$

$$g_1(3/1)=f(3,1)/f_2(1)=(0.05)/(0.45)=0.11$$

وبنفس الطريقة لحساب دالة التوزيع الاحتمالي الشرطية إلى Y عندما يكون $X=2$ وهي:

$$g_2(0/2)=f(2,0)/f_1(2)=(0.15)/(0.30)=0.50$$

$$g_2(1/2)=f(2,1)/f_1(2)=(0.10)/(0.30)=0.33$$

$$g_2(2/2)=f(2,2)/f_1(2)=(0.05)/(0.30)=0.17$$

ولحساب دالة التوزيع التجميعية الشرطية كما يأتي:

لتكن $G_2(y/x)$ تمثل دالة التوزيع التجميعية الشرطية إلى Y عندما $X=2$ معطى فنحصل على:

$$G_2(0/2)=g_2(0/2)=0.50$$

$$G_2(1/2)=g_2(0/2)+g_2(1/2)=0.50+0.33=0.83$$

$$G_2(2/2)=g_2(0/2)+g_2(1/2)+g_2(2/2)=0.50+0.33+0.17=1.00$$

3-4-7 دالة التوزيع الاحتمالي الشرطية للمتغيرات العشوائية المستمرة:

Conditional Probability Distribution Function Of Continuous Random Variables:

تعريف 7-8

ليكن X و Y متغيرين عشوائيين مستمرين لهما دالة التوزيع الاحتمالي المشتركة (joint $p.d.f$) $f(x,y)$ ودالتي التوزيع الاحتمالي الهامشية (marginal $p.d.f$) $f_1(x)$ و $f_2(y)$ على التوالي حيث كلاهما تمثلان دوال موجبة فان :

$$g_1(x/y)=\frac{f(x,y)}{f_2(y)} \qquad \text{(7-13)}$$

حيث $f_2(y)>0$.

فإن $g_1(x/y)$ تسمى دالة التوزيع الاحتمالي الشرطية للمتغير العشوائي المستمر X عندما يكون المتغير العشوائي المستمر $Y=y$ معطى.

وبنفس الأسلوب نحصل على دالة التوزيع الاحتمالي الشرطية للمتغير العشوائي المستمر Y عندما $X=x$ معطى وهي:

$$g_2(y/x)=\frac{f(x,y)}{f_1(x)} \qquad \text{(7-14)}$$

حيث $f_1(x)>0$.

نلاحظ بان كل من الدالتين $g_1(x/y)$ و $g_2(y/x)$ تحققان شروط دوال التوزيع الاحتمالي المستمر وهي:

- $g_1(x/y)\geq 0$ و $g_2(y/x)\geq 0$.

○ $\int_{-\infty}^{\infty} g_1(x/y)dx = \int_{-\infty}^{\infty} \frac{f(x,y)dx}{f_2(y)} = \frac{f_2(y)}{f_2(y)} = 1$ كذلك فان :

$$\int_{-\infty}^{\infty} g_2(y/x)dy = \int_{-\infty}^{\infty} \frac{f(x,y)dy}{f_1(x)} = \frac{f_1(x)}{f_1(x)} = 1$$

ملاحظة:

نلاحظ بان $\int_a^b g_1(x/y)dx$ يمثل الاحتمال $P(a \le X \le b / Y = y)$ وبما أن $P(Y = y) = 0$ في حالة المتغير العشوائي المستمر فأنه لايمكن لنا أن نستخدم طريقة التعريف السابقة في تعريف دالة الاحتمال الشرطي،ولكن على أية حال سنستخدم التكامل لتعريف هذه الاحتمالية.

مثال 7-10

ليكن (X,Y) متغيرين عشوائيين مستمرين لهما دالة التوزيع الاحتمالي المشتركة الآتية:

$$f(x,y) = \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} exp\left\{-\frac{1}{2}\frac{(x^2 - 2\rho xy + y^2)}{1-\rho^2}\right\}$$

حيث أن $-\infty < x, y < \infty, |\rho| < 1$.

الحل

بما أن دالة التوزيع الاحتمالي الهامشية إلى X هي:

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)dy = (2\pi)^{-\frac{1}{2}} exp(-\frac{1}{2}x^2) \quad , -\infty < x < \infty$$

فان دالة التوزيع الاحتمالي الشرطية إلى Y عندما $X = x$ معطى هي:

$$g(y/x) = \frac{f(x,y)}{f_1(x)} = \frac{(2\pi)^{-1}(1-\rho^2)^{-\frac{1}{2}} e^{-\frac{1}{2}(x^2 - 2\rho xy + y^2)/91-\rho^2)}}{(2\pi)^{-\frac{1}{2}} e^{-x^2/2}}$$

$$= \frac{1}{\{2\pi(1+\rho^2)\}^{1/2}} e^{-\frac{1}{2}(y-\rho x)^2/(1-\rho^2)} \quad , -\infty < y < \infty$$

وهذا يمثل توزيعاً طبيعياً بوسط مقداره (ρx) وتباين $(1-\rho^2)$.

5-7 المتغيرات العشوائية المستقلة Independent Random Variables

سبق وان ناقشنا فكرة الأحداث المستقلة في الفصل الثاني أثبتنا بان الحدثين A و B مستقلين إذا كان $P(A \cap B) = P(A)P(B)$. وهنا سوف نتطرق إلى الاستقلالية في المتغيرات العشوائية، والتي سيكون لها دوراً مهما في دراساتنا المتقدمة وخاصة في مجال نظرية العينات.[2]

تعريف 7-9

ليكن كل من X و Y متغيرين عشوائيين . نقول بأنهما مستقلان (independent) إذا كان :

$$F(x,y) = F_1(x)F_2(y) \quad \text{.......................................(7-15)}$$

لكل قيم (x,y).

ولكي نبين بأن المتغيرين العشوائيين X و Y مستقلان ،نطبق النظرية الآتية والتي تسمح لنا باستخدام دالة التوزيع الاحتمالي المشتركة بدلاً من دالة التوزيع التجميعية الواردة في التعريف (7-9) .

نظريه 7-2

لتكن $f(x,y)$ دالة التوزيع الاحتمالي المشتركة للمتغيرين العشوائيين X و Y ولتكن $f_1(x)$ و $f_2(y)$ هما دالتين التوزيع الهامشي لكلٍ من X و Y على التوالي . فأن X و Y متغيرين عشوائيين مستقلين اذا وإذا فقط كان:

$$f(x,y) = f_1(x)f_2(y) \quad \text{.......................................(7-16)}$$

لكل قيم x و y.

البرهان

ليكن X و Y متغيرين عشوائيين مستمرين مستقلين فان:

$$f(x,y)=\frac{\partial^2 F(x,y)}{\partial x\,\partial y}=\frac{\partial^2\{F_1(x)F_2(y)\}}{\partial x\,\partial y}$$

$$=\frac{\partial F_1(x)}{\partial x}\frac{\partial F_2(y)}{\partial y}=f_1(x)f_2(y)$$

ولبرهان الاتجاه الآخر من النظرية نفرض بان العلاقة $(7-16)$ متحققة فان:

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\,dv\,du=\int_{-\infty}^{x}\int_{-\infty}^{y}f_1(u)f_2(v)\,dv\,du$$

$$=\left\{\int_{-\infty}^{x}f(u)\,du\right\}\left\{\int_{-\infty}^{y}f(v)\,dv\right\}=F_1(x)F_2(y)$$

وهذا يعني بان X و Y متغيران عشوائيان مستقلان.

ملاحظة:

يمكن أن نبرهن بان المتغيرين العشوائيين X و Y مستقلين باستخدام دالة التوزيع الاحتمالي الشرطية وكما يأتي:

إذا كانت كل من $g_1(x/y)$ و $g_2(y/x)$ تمثلان دالة التوزيع الاحتمالي الشرطية إلى X عندما $Y=y$ معطى و Y عندما $X=x$ معطى وذلك على التوالي.

وبما أن تعريف دالة التوزيع الاحتمالي الشرطية هو:

$$g_2(x/y)=\frac{f(x,y)}{f_1(x)} \quad \text{و} \quad g_1(x/y)=\frac{f(x,y)}{f_2(y)}$$

فإن:

$$f(x,y)=f_2(y)g_1(x/y) \quad \text{و} \quad f(x,y)=f_1(x)g_2(y/x) \quad \ldots\ldots\ldots\textbf{(7-17)}$$

وبمقارنة العلاقة $(7-16)$ و $(7-17)$ نحصل على أن X و Y مستقلين اذا وإذا فقط كان:

$g_1(x/y)=f_1(x)$ و $g_2(y/x)=f_2(x)$ لكل القيم الحقيقية إلى x و y.

وهذا يعني بان لكل متغير عشوائي فان التوزيع الهامشي والشرطي يجب أن يكونان متكافئان.

مثال11-7

لتكن دالة التوزيع التجميعية للمتغيرين العشوائيين المستمرين X و Y هي:

$$F(x,y)=\begin{cases}1-e^{-x}-e^{-y}+e^{-(x+y)} & ,x,y\geq 0\\ 0 & ,elsewhere\end{cases}$$

- أحسب دالة التوزيع الاحتمالي الهامشية للمتغيرين X و Y؟
- أحسب دالة التوزيع التجميعية للمتغيرين X و Y؟
- بين فيما اذا كان كل من X و Y مستقلين ام لا؟

الحل

من الدالة $F(x,y)$ نحصل على دالة التوزيع الاحتمالية المشتركة وكما يأتي:

$$\frac{\partial^2 F(x,y)}{\partial x\,\partial y}=f(x,y)=\begin{cases}e^{-(x+y)} & ,x,y\geq 0\\ 0 & ,elsewhere\end{cases}$$

وبذلك فان دالة التوزيع الاحتمالي الهامشية تكون:

$$f_1(x)=\int_{-\infty}^{\infty}f(x,y)dy=\begin{cases}e^{-x} & ,x>0\\ 0 & ,elsewhere\end{cases}$$

و

$$f_2(y)=\int_{-\infty}^{\infty}f(x,y)dx=\begin{cases}e^{-y} & ,y>0\\ 0 & ,elsewhere\end{cases}$$

أما بالنسبة إلى دالة التوزيع التجميعية الهامشية فهي:

$$F_1(x) = F(x, \infty) = \begin{cases} 0 & , x < 0 \\ 1 - e^{-x} & , x \geq 0 \end{cases}$$

و

$$F_2(y) = F(y, \infty) = \begin{cases} 0 & , y < 0 \\ 1 - e^{-y} & , y \geq 0 \end{cases}$$

وبما أن $F(x, y) = F_1(x) F_2(y)$ لجميع قيم x و y، فان المتغيرين العشوائيين X و Y مستقلان (independent). وهذا يؤكد بان :

$f(x, y) = f_1(x) f_2(y)$ لجميع قيم x و y .

مثال 7-12

صندوق يحتوي على (20) شمعة عيد ميلاد (5) منها حمراء ، (10) زرقاء و(5) بيضاء. سحبت منها عينه تتكون من(5) شموع عشوائياً ، وليكن X يمثل عدد الشموع الحمراء و Y يمثل عدد الشموع الزرقاء المسحوبة .

- أحسب دالة التوزيع الاحتمالي المشتركة بالنسبة إلى X و Y ؟
- أحسب دالة التوزيع الاحتمالي الهامشية بالنسبة إلى Y ؟
- أحسب التوزيع الشرطي إلى X عندما $Y = y$ معطى ، أستخدم العلاقة الآتية لحساب النتائج :

$\sum_{all} \binom{a}{x}\binom{b}{n-x} = \binom{a+b}{n}$ حيث أن $\binom{a}{n} = 0$ لجميع قيم $x < a$ و a, b, n قيم صحيحة موجبة وأن $n \leq a + b$.

الحل

عدد طرق اختيار عينة من الحجم (5) من (20) شمعة $= \binom{20}{5}$.

وإن جميع هذهِ الاختيارات العشوائية تكون متماثلة .

نفرض بأن العينة x هي شمعه حمراء ، y شمعه زرقاء فإن $5-x-y$ شمعةً بيضاء .

عدد طرق اختيار x شمعة حمراء من العينة $=\binom{5}{x}$.

عدد طرق اختيار y شمعة زرقاء من العينة $=\binom{10}{y}$.

عدد طرق اختيار $5-x-y$ شمعةً بيضاء من العينة $=\binom{5}{5-x-y}$.

ولأن اختيار الشموع الحمراء مرتبط مع كل اختيار لشمعة من الشموع الزرقاء والتي جميعها ترتبط مع كل اختيار من الشموع البيضاء فإن العدد الكلي للعينات الممكنة التي تحتوي شموعاً حمراء و زرقاء وبيضاء هو : $\binom{5}{x}\binom{10}{y}\binom{5}{5-x-y}$.

لذلك فإن دالة التوزيع الاحتمالي المشتركة إلى X و Y هي :

$$f(x,y)=P(X=x,Y=y)=\frac{\binom{5}{x}\binom{10}{y}\binom{5}{5-x-y}}{\binom{20}{5}}$$

حيث أن $y\geq 0$ ، $x\geq 0$ ، $0\leq x+y\leq 5$.

ودالة التوزيع الاحتمالي الهامشي بالنسبة إلى Y هي :

$$f_2(y)=\sum_{x=0}^{5-y}\frac{\binom{5}{x}\binom{10}{y}\binom{5}{5-x-y}}{\binom{20}{5}}$$

حيث أن $0\leq y\leq 5$.

ودالة التوزيع الاحتمالي الهامشي بالنسبة إلى X هي:

$$f_1(x)=\sum_{y=0}^{5-x}\frac{\binom{5}{x}\binom{15}{5-x}}{\binom{20}{5}}$$

حيث أن $0\le x\le 5$.

ولحساب دالة التوزيع الشرطية بالنسبة إلى X عندما $Y=y$ معطى فتكون:

$$g_1(x/y)=\frac{f(x,y)}{f_2(y)}=\frac{\binom{5}{x}\binom{5}{5-x-y}}{\binom{10}{5-y}}$$

حيث أن $0\le x\le 5-y, 0\le y\le 5$.

تمرين:

اثبت من خلال مثال(8-7)بان X و Y متغيرين عشوائيين مستقلين ؟

7-6 التوقع للدوال ذات المتغيرين العشوائيين

Expectations Of Functions Of Pairs Of Random Variables

في هذه الوحدة سنوسع مفهوم التوقع لكي يشمل التوقع للمتغيرين العشوائيين،وسنتناوله من خلال بعض الأمثلة الآتية:

مثال7-13

رمي حجر نرد مرتين،فإذا كانت دالة التوزيع الاحتمالي المشتركة للمتغيرين X و Y تعتمد على نتائج الرميات ،حيث انه اذا كان كلا العددين زوجيين وغير متساويين أو فرديين وغير متساويين فإننا نحصل على مجموع هذين العددين،أما اذا كان احد الأعداد فردي والآخر زوجي فنحصل على مجموع سالب للعددين،وإذا كان العددان فرديين ومتساويين فنحصل على القيمة المشتركة التي تمثل العددين ،

وإذا كان العددان زوجيين ومتساويين فنحصل على القيمة المشتركة السالبة لهما.وبذلك يمكن تكوين دالة التوزيع الاحتمالي المشتركة وكما يأتي:

ليكن كل من X و Y متغيرين عشوائيين الأول يمثل نتائج الرمية الأولى لحجر النرد والثاني يمثل نتائج الرمية الثانية .اذا فرضنا بان U يمثل دالة لهذين المتغيرين العشوائيين X و Y ونرمز لها بالرمز $U(X,Y)$ فإنها تعرف بالشكل الآتي:

$$U(X,Y)=\begin{cases} X+Y \quad if \quad X\neq Y and \quad X,Y are \quad both \quad even \quad or \quad odd \\ -(X+Y) \quad if \quad X \quad is \quad odd \quad and \quad Y \quad is \quad even \quad or \quad op. \\ X \quad if \quad X=Y \quad and \quad X \quad is \quad odd \\ -X \quad if \quad X=Y \quad and \quad X \quad is \quad even \end{cases}$$

وتكون القيم الممكنة إلى (x,y) بالنسبة إلى (X,Y) كما في الجدول أدناه:

x \ y	1	2	3	4	5	6
1	1	-3	4	-5	6	-7
2	-3	-2	-5	6	-7	8
3	4	-5	3	-7	8	-9
4	-5	6	-7	-4	-9	10
5	6	-7	8	-9	5	-11
6	-7	8	-9	10	-11	-6

وإذا اعتبرنا U متغير عشوائي بحد ذاته فان التوزيع الاحتمالي له نحصل علية من الجدول أعلاه، فمثلا هناك (4)قيم من(36)قيمة تحمل الرقم (-5)

فإن $P(U=-5)=\frac{4}{36}$ وعلية يمكن إيجاد جميع قيم الاحتمالات الممكنة والتي تكون متماثلة باحتمال $\frac{1}{36}$(equally likely)وكما يأتي:

$U:$	-11	-9	-7	-6	-5	-4	-3	-2	1	3	4	5	6	8	10
p $(U=u)$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{6}{36}$	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{4}{36}$	$\frac{2}{36}$

وهناك طريقتان لحساب التوقع بالنسبة للمتغير العشوائي U وكما يلي:

$$E(U)=\sum_{all\ \ u} UP(U=u)=-11(\frac{2}{36})+(-9)(\frac{4}{36})+\ldots+10(\frac{2}{36})=-\frac{5}{4}$$

أو نتعامل مع U كدالة إلى X و Y فنحصل على:

$$E(U)=1(\frac{1}{36})+(-3)(\frac{1}{36})+\ldots+(-6)(\frac{1}{36})=-\frac{5}{4}$$

وان الفائدة من الطريقة الثانية تجعلنا نستطيع أن نتعامل مع دالة التوزيع الاحتمالي المشتركة إلى X و Y دون الحاجة لاشتقاق توزيع احتمالي بالنسبة إلى U، والنظرية أدناه توضح ذلك.

نظرية 7-3

ليكن (X,Y) متغير عشوائي ثنائي له دالة التوزيع الاحتمالي المشتركة $f(x,y)$ إذا كان $U=u(X,Y)$ دالة إلى (X,Y)، فان التوقع (expected) للقيمة U، إذا وجد فيعطى بالصيغة الآتية:

$$E(U)=E\{u(X,Y)\}=\sum_{(x,y)\in R_{X,Y}}\sum u(x,y)f(x,y) \quad \ldots\ldots(7\text{-}18)$$

في حالة المتغير العشوائي المتقطع.

أما في حالة المتغير العشوائي المستمر فيعطى بالصيغة الآتية:

$$E(U)=E\{u(X,Y)\}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}u(x,y)f(x,y)dxdy \quad \ldots\ldots\ldots\ldots(7\text{-}19)$$

نظرية 7-4

ليكن (X,Y) متغير عشوائي ثنائي له دالة الكثافة الاحتمالية المشتركة $f(x,y)$ وان $U=a_1X+a_2Y$ فان التوقع الى U:

$$E(U)=a_1E(X)+a_2E(Y) \qquad \text{.................................(7-20)}$$

والتباين الى U هو:

$$var(U)=a_1^2\ \ var(X)+a_2^2\ \ var(Y)+2a_1a_2\,cov(X,Y) \qquad \text{......(7-21)}$$

حيث أن $cov(X,Y)=E(XY)-E(X)E(Y)$ يسمى التغاير(covariance) للمتغيرين X و Y.

البرهان

لبرهان العلاقة (20 – 7) نضع :

$g(x,y)=a_1x+a_2y$ وباستخدام العلاقة (19 – 7) فان:

$$E(U)=E(a_1X+a_2Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a_1x+a_2y)f(x,y)dxdy$$

$$=a_1\int_{-\infty}^{\infty}x\int_{-\infty}^{\infty}f(x,y)dydx+a_2\int_{-\infty}^{\infty}y\int_{-\infty}^{\infty}f(x,y)dxdy$$

وبما أن $\int_{-\infty}^{\infty}f(x,y)dy=f_1(x)$ و $\int_{-\infty}^{\infty}f(x,y)dx=f_2(y)$ فان:

$E(U)=a_1\int_{-\infty}^{\infty}xf_1(x)+a_2\int_{-\infty}^{\infty}yf_2(y)$ حيث أن $f_1(x),f_2(y)$ هما الدوال الهامشية للمتغيرين العشوائيين X و Y، وبما أن $\int_{-\infty}^{\infty}xf_1(x)dx=E(X)$ وان $\int_{-\infty}^{\infty}yf_2(y)dy=E(Y)$ ، وبالتعويض عن هذين التكاملين أعلاه نحصل على:

$$E(U)=a_1E(X)+a_2E(Y)$$

ولبرهان العلاقة(7-21) أعلاه نضع:

$$U^2=(a_1X+a_2Y)^2=a_1{}^2X^2+a_2{}^2Y^2+2a_1a_2XY$$

وباستخدام العلاقة (7-19) ووضع:

$$g(x,y)==a_1{}^2x^2+a_2{}^2y^2+2a_1a_2xy$$

فإننا نحصل على:

$$E(U^2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a_1^2x^2 + a_2^2y^2 + 2a_1a_2xy)f(x,y)dxdy$$

$$= a_1^2\int_{-\infty}^{\infty}x^2\int_{-\infty}^{\infty}f(x,y)dydx + a_2^2\int_{-\infty}^{\infty}y^2\int_{-\infty}^{\infty}f(x,y)dxdy$$

$$+2a_1a_2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)dxdy$$

$$= a_1^2\int x^2 f_1(x)dx + a_2^2\int y^2 f_2(y)dy + 2a_1a_2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)dxdy$$

$$= a_1^2E(X^2) + a_2^2E(Y^2) + 2a_1a_2E(XY)$$

وبما أن :

$$var(U) = E(U^2) - \{E(U)\}^2$$

$$= a_1^2\left[E(X^2) - \{E(X)\}^2\right] + a_2^2\left[E(Y^2) - \{E(Y)\}^2\right]$$

$$+2a_1a_2\left[E(XY) - E(X)E(Y)\right]$$

$$var(U) = a_1^2 \ var(X) + a_2^2 \ var(Y) + 2a_1a_2cov(X,Y)$$

ملاحظة:

هناك حالتان مهمتان يجب الإشارة إليهما:

1. اذا كانت $a_1 = a_2 = 1$ فان $E(X+Y) = E(X) + E(Y)$ و...

$$var(X+Y) = var(X) + var(Y) + 2cov(X,Y) \quad ....(7\text{-}22)$$

2. إذا كانت $a_1 = 1, a_2 = -1$ فان $E(X-Y) = E(X) - E(Y)$ و...

$$var(X-Y) = var(X) + var(Y) - 2cov(X,Y) \quad ....(7\text{-}23)$$

مثال 7-14

نظام الكتروني يتكون من عدة مكونـات ،وهـذه المكونـات تخـضع للعطـل عنـد العمل لذلك ترتبط بمكونات احتياطية تحل محلها حالاً في حالـة تعرضـها للعطـل

وتعمل كبديل عنها مباشرةً ، ولهذا يمكن أن نمثل استمرارية العمل للنظام(بقاء النظام) بالمتغير $T = X + Y$،حيث أن X يمثل مكونات النظام و Y يمثل المكونات الاحتياطية .اذا كان X و Y متغيرين عشوائيين مستمرين ومستقلين ولهما دالتي التوزيع الاحتمالي $\lambda_1 e^{-\lambda_1 x}$ و $\lambda_1 e^{-\lambda_1 x}$ على التوالي حيث أن $0 < x , y < \infty$ ، أحسب متوسط البقاء والتباين للنظام؟

الحل

نلاحظ بان كل من الدوال الهامشية $f_1(x) = \lambda_1 e^{-\lambda_1 x}$ و $f_2(y) = \lambda_2 e^{-\lambda_2 y}$ هي دوال أُسية بمعالم λ_1 و λ_2 حيث سبق وان حسبنا:

$E(X) = \dfrac{1}{\lambda_1}$ و $E(Y) = \dfrac{1}{\lambda_2}$ وكذلك $var(X) = \dfrac{1}{\lambda_1^2}$ و $var(Y) = \dfrac{1}{\lambda_2^2}$.

إذن نحصل على متوسط بقاء النظام :

$E(T) = E(X) + E(Y) = \dfrac{1}{\lambda_1} + \dfrac{1}{\lambda_2}$،وبما أن المتغيرين X و Y مستقلين فان دالة التوزيع التجميعية المشتركة لهما هي:

$$f(x, y) = f_1(x) f_2(y) = (\lambda_1 e^{-\lambda_1 x})(\lambda_2 e^{-\lambda_2 y}) = \lambda_1 \lambda_2 e^{-(\lambda_1 x + \lambda_2 y)}, x, y > 0$$

كذلك فان:

$$E(XY) = \int_0^\infty \int_0^\infty xy\, \lambda_1 \lambda_2 e^{-(\lambda_1 x + \lambda_2 y)} dx dy$$
$$= \left\{ \int_0^\infty x\, \lambda_1 e^{-\lambda_1 x} dx \right\} \left\{ \int_0^\infty y\, \lambda_2 e^{-\lambda_2 y} dy \right\} = \frac{1}{\lambda_1} \frac{1}{\lambda_2}$$

وبما أن $cov(X, Y) = E(XY) - E(X)E(Y)$

فانه بالتعويض عن قيم $E(XY)$ ، $E(X)$ و $E(Y)$ نحصل على أن:

$cov(X, Y) = 0$.

وعلية يكون حساب التباين للنظام هو:

$$var(T) = var(X) + var(Y) + 2\, cov(X, Y)$$

$$= \frac{1}{\lambda_1^{\,2}} + \frac{1}{\lambda_2^{\,2}} + 0 = \frac{1}{\lambda_1^{\,2}} + \frac{1}{\lambda_2^{\,2}}$$

من خلال المثال أعلاه يمكن وضع الاستنتاج الآتي بصيغة نظرية.

نظرية 7-5

ليكن كل من X و Y متغيرين عشوائيين لهما دالة التوزيع المشتركة $f(x,y)$. اذا كان X و Y مستقلين فان:

$$E(XY) = E(X)E(Y)$$

البرهان

اذا كان X و Y متغيرين عشوائيين مستمرين فان:

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_1(x)f_2(y)\,dx\,dy$$

$$= \left\{\int_{-\infty}^{\infty} xf_1(x)\,dx\right\}\left\{\int_{-\infty}^{\infty} yf_2(y)\,dy\right\} = E(X)E(Y)$$

ويمكن البرهان في حالة المتغيرين العشوائيين X و Y متقطعين بالأسلوب نفسه.

ملاحظة:

اذا كان X و Y مستقلين فان $\operatorname{cov}(X,Y)=0$ ، لذلك نستنتج من العلاقتين $(7-22)$ و $(7-23)$ بان:

$$var(X+Y) = var(X-Y) = var(X) + var(Y) \quad \ldots\ldots\ldots\ldots\ldots (7\text{-}24)$$

7-7 العزوم الثنائية Bivariate Moments

سوف نوضح في هذه الفقرة مفهوم العزم الأحادي لمتغير واحد إلى العزوم الثنائية ودوالها المولدة للعزوم،حيث أن هذه الدوال تعطينا فكرة جيدة عن بعض الخواص المهمة للتوزيعات المرتبطة بها ،وسوف ندرس إحدى هذه الخواص المهمة وهي مفهوم الارتباط بين المتغيرين العشوائيين.

تعريف 7-10

ليكن (X,Y) متغير عشوائي ثنائي له دالة التوزيع الاحتمالي المشتركة $f(x,y)$ ، فان العزم الثنائي (bivariate moment) من الرتبة (r,s)

للمتغير (X,Y) حول نقطة الأصل ، اذا وجد،فيعرف كما يأتي:[2]

$$E(X^rY^s) = \sum_{(x,y)\in R_{X,Y}} \sum x^r y^s f(x,y) \qquad \text{(7-25)}$$

في حالة كون المتغير العشوائي الثنائي (X,Y) من النوع المتقطع.

أما في حالة كون المتغير العشوائي (X,Y) من النوع المستمر فان العزم الثنائي يعرف كمايلي:

$$E(X^rY^s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^r y^s f(x,y)\,dxdy \qquad \text{(7-26)}$$

إن العزم الثنائي حول نقطة الأصل $E(X^r,Y^s)$ سوف نرمز له بالرمز $\mu'_{r,s}$ وهو يشير إلى العزم المختلط للتوزيع(mixed moment) من الرتبة (r,s) .

اذا كان $s=0$ فان $E(X^r)=\mu'_{r,0}$ يسمى العزم r حول نقطة الأصل للتوزيع الهامشي X .

وإذا كان $r=0$ فان $E(Y^s)=\mu'_{0,s}$ يسمى العزم s حول نقطة الأصل للتوزيع الهامشي Y .

وبشكل محدد يمكن أن نعبر عن $\mu'_{1,0}$ و $\mu'_{0,1}$ بأنهما يمثلان المتوسطان للتوزيعات الهامشية بالنسبة إلى X و Y على التوالي.

وكما لاحظنا في حالة المتغير العشوائي الواحد يكون من المفيد العمل حول المتوسطات، ولهذا سوف نستخدم هذه المتوسطات لتعريف العزم المركزي الثنائي بالشكل الآتي:

$$\mu_{r,s} = E\left[\{X-E(X)\}^r\{Y-E(Y)\}^s\right] \qquad \text{(7-27)}$$

$$\mu_{r,s} = \sum_{(x,y)\in R_{X,Y}} \sum (x-\mu'_{1,0})^r (y-\mu'_{0,1})^s f(x,y) \text{ ....(7-28)}$$

في حالة المتغير العشوائي الثنائي المتقطع.

أما في حالة المتغير العشوائي الثنائي المستمر فيعرف بالشكل الآتي:

$$\mu_{r,s} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu'_{1,0})^r (y-\mu'_{0,1})^s f(x,y)\,dxdy \text{ ....(7-29)}$$

عندما $r=s=1$ فإننا نحصل على:

$$\mu_{1,1} = E\left[\{X-E(X)\}\{Y-E(Y)\}\right]$$
$$= E(XY) - E\{Y\,E(X)\} - E\{X\,E(Y)\} + E\{E(X)E(Y)\}$$

وبما أن كل من $\mathrm{E}(X)$ و $\mathrm{E}(Y)$ ثوابت فان:

$$\mu_{1,1} = E(XY) - E(Y)E(Y) - E(X)E(Y) + E(X)E(Y)$$
$$\mu_{1,1} = E(XY) - E(Y)E(Y))$$
$$= cov(X,Y) \text{ ....(7-30)}$$

وهذا يمثل التغاير (covariance) بين X و Y.

ويمكن لنا أن نعبر عن العزم المركزي $\mu_{r,s}$ باستخدام العزوم الثنائية حول نقطة الأصل $\mu'_{r,s}$. وعليه سوف نستخدم التوسيع الثنائي الحدين في العلاقة $(7-29)$ للحصول على العلاقة الآتية:

$$\mu_{r,s} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{\sum_{i=0}^{r}\binom{r}{i}(-1)^i x^{r-i}\mu'^{\,i}_{1,0}\right\}\left\{\sum_{j=0}^{s}\binom{s}{j}(-1)^j y^{s-j}\mu'^{\,j}_{0,1}\right\} f(x,y)\,dxdy$$
$$= \sum_{i=0}^{r}\sum_{j=0}^{s}\binom{r}{i}\binom{s}{j}(-1)^{i+j}\mu'^{\,i}_{1,0}\mu'^{\,j}_{0,1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^{r-i}y^{s-j} f(x,y)\,dxdy$$
$$= \sum_{i=0}^{r}\sum_{j=0}^{s}\binom{r}{i}\binom{s}{j}(-1)^{i+j}\mu'^{\,i}_{1,0}\mu'^{\,j}_{0,1}\mu'_{r-i,s-j} \text{ ....(7-31)}$$

وعلى سبيل المثال اذا فرضنا أن $r=s=1$ سنحصل على:

$\mu_{1,1} = \mu'_{1,1} - \mu'_{1,0}\mu'_{0,1}$ وهكذا اذا فرضنا أن $r=2, s=1$ فسنحصل على:

$$\mu_{2,1} = \mu'_{2,1} - \mu'_{2,0}\mu'_{0,1} - 2\,\mu'_{1,1}\mu'_{1,0} + 2\,\mu'^{2}_{1,0}\mu'_{0,1}$$

تعريف 7-11

اذا كان كل من X و Y متغيرين عشوائيين لهما دالة التوزيع الاحتمالي المشتركة $f(x,y)$ فان دالة العزم المولدة المشتركة (Joint Moment Generating Function) لهما في حالة كونهما مستمرين هي:

$$M_{X,Y}(t_1,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1x+t_2y} f(x,y)\,dxdy \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(7\text{-}32)$$

حيث أن $|t_1| < h_1$ و $|t_2| < h_2$، وان h_1 و h_2 قيم ثابتة موجبة.

أما في حالة كون المتغيرين العشوائيين X و Y متقطعين فان دالة العزم المولدة المشتركة لهما هي:

$$M_{X,Y}(t_1,t_2) = \sum_{(x,y)\in R_{X,Y}} \sum e^{t_1x+t_2y} f(x,y) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(7\text{-}33)$$

حيث أن $|t_1| < h_1$ و $|t_2| < h_2$، وان h_1 و h_2 قيم ثابتة موجبة.

إن التعريف أعلاه يبين لنا بان $M_{X,Y}(t_1,t_2) = E(e^{t_1x+t_2y})$ إذا كان التوقع موجوداً. وعليه يمكن أن نبين بان دالتي العزم الهامشية المولدة هما:

$M_{X,Y}(t_1,0) = M_X(t_1)$ و $M_{X,Y}(0,t_2) = M_Y(t_2)$ للمتغيرين X و Y على التوالي. وكذلك يمكن لنا أن نلاحظ بأنه من خلال استخدام الاشتقاق الجزئي لجميع الرتب عند $t_1 = t_2 = 0$ في حالة وجود دالة العزم $M_{X,Y}(t_1,t_2)$ وكما يأتي:

$$\frac{\partial^{r+s} M_{X,Y}(t_1,t_2)}{\partial t_1^{\,r}\partial t_2^{\,s}} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^r y^s e^{t_1x+t_2y} f(x,y)\,dxdy$$

في حالة $t_1 = t_2 = 0$ فان $e^{t_1x+t_2y} = 1$ وعليه يكون:

$$\left\{ \frac{\partial^{r+s} M_{X,Y}(t_1,t_2)}{\partial t_1^{\,r}\partial t_2^{\,s}} \right\}_{t_1=t_2=0} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^r y^s f(x,y)\,dxdy = \mu'_{r,s} \quad \ldots(7\text{-}34)$$

في حالة المتغير العشوائي المستمر.

وبالأسلوب نفسه نستنتج بأنه في حالة المتغير العشوائي المتقطع فان:

$$\left\{ \frac{\partial^{r+s} M_{X,Y}(t_1,t_2)}{\partial t_1^{\,r} \partial t_2^{\,s}} \right\}_{t_1=t_2=0} = \sum_{(x,y)\in R_{X,Y}} \sum x^r y^s f(x,y) = \mu'_{r,s} \quad \ldots(7\text{-}35)$$

وفيما يأتي مثال على متغيرين عشوائيين مستمرين.

مثال 7-15

ليكن X و Y متغيرين عشوائيين مستمرين لهما دالة التوزيع الاحتمالي المشتركة الآتية:

$$f(x,y) = \begin{cases} e^{-y}, 0 < x < y < \infty \\ 0, \qquad elsewher \end{cases}$$

إحسب ما يأتي:

- دالة العزم المولدة المشتركة بالنسبة إلى X و Y ؟
- دالتي العزم المولدة الهامشية؟
- التغاير والتباين لكل من X و Y ؟

الحل

دالة العزم المولدة المشتركة تحسب كما يأتي:

$$M_{X,Y}(t_1,t_2) = \int_0^\infty \int_x^\infty e^{(t_1x+t_2y)} e^{-y} dydx = \int_0^\infty \int_x^\infty e^{(t_1x+t_2y-y)} dydx$$

$$= \int_0^\infty \left\{ \frac{1}{(t_2-1)} e^{(t_1x+t_2y)} \right\}_x^\infty dx$$

$$= \int_0^\infty \left(\left\{ \frac{1}{(t_2-1)} e^{(t_1x+t_2\infty-\infty)} \right\} - \left\{ \frac{1}{(t_2-1)} e^{(t_1x+t_2x-x)} \right\} \right) dx$$

$$= \int_0^\infty \frac{1}{(1-t_2)} e^{x(t_1+t_2-1)} dx = \int_0^\infty \frac{e^{-x(1-t_1-t_2)}}{1-t_2} dx$$

$$= \frac{e^{-x(1-t_1-t_2)}}{-(1-t_2)(1-t_1-t_2)} \Bigg|_0^\infty$$

وبعد إجراء التعويض بحدود التكامل نحصل على:

$$M_{X,Y}(t_1,t_2)=\frac{1}{(1-t_2)(1-t_1-t_2)}$$

وهي تمثل دالة العزم المولدة المشتركة.

ولحساب دالة العزم المولدة الهامشية فان من الدالة اعلاه يكون:

$M_X(t_1)=M_{X,Y}(t_1,0)=\frac{1}{(1-t_1)}\quad ,t_1<1$ دالة العزم الهامشية إلى X

$M_Y(t_2)=M_{X,Y}(0,t_2)=\frac{1}{(1-t_2)^2}\quad ,t_2<1$ دالة العزم الهامشية إلى Y وبأخذ المشتقة الأولى والثانية بالنسبة إلى t_1 و t_2 للدالتين أعلاه نحصل:

$$\frac{\partial^2 M_X(t_1)}{\partial t_1^{\,2}}=\frac{2}{(1-t_1)^3}\quad \text{و} \quad \frac{\partial M_X(t_1)}{\partial t_1}=\frac{1}{(1-t_1)^2}$$

$$\frac{\partial^2 M_Y(t_2)}{\partial t_2^{\,2}}=\frac{6}{(1-t_2)^4}\quad \text{و} \quad \frac{\partial M_Y(t_2)}{\partial t_2}=\frac{2}{(1-t_2)^3}$$

كذلك نأخذ المشتقة الثانية لدالة العزم المولدة المشتركة أي أن :

$$\frac{\partial^2 M_{X,Y}(t_1,t_2)}{\partial t_1\partial t_2}=\frac{3-3t_2-t_1}{(1-t_2)^2(1-t_1-t_2)^3}$$

وبعد التعويض في قيم المشتقات أعلاه عن قيم t_1 و t_2 كل حسب مشتقتها نحصل على ما يأتي:

$$\frac{\partial M_X(t_1)}{\partial t_1}=\mu'_{1,0}=\frac{1}{(1-t_1)^2}=\frac{1}{(1-0)^2}=1=E(X)$$

$$\frac{\partial M_Y(t_2)}{\partial t_2}=\mu'_{0,1}=\frac{2}{(1-t_2)^3}=\frac{2}{(1-0)^3}=2=E(Y)$$

$$\frac{\partial^2 M_X(t_1)}{\partial t_1^{\,2}}=\mu'_{2,0}=\frac{2}{(1-t_1)^3}=\frac{2}{(1-0)^3}=2=E(X^2)$$

$$\frac{\partial^2 M_Y(t_2)}{\partial t_2^{\,2}}=\mu'_{0,2}=\frac{6}{(1-t_2)^4}=\frac{6}{(1-0)^4}=6=E(Y^2)$$

$$\frac{\partial^2 M_{X,Y}(t_1,t_2)}{\partial t_1 \partial t_2} = \mu'_{1,1} = \frac{3-3t_2-t_1}{(1-t_2)^2(1-t_1-t_2)^3}\Bigg|_{t_1=t_2=0}$$

$$= \frac{3-0-0}{(1-0)^2(1-0-0)^3} = 3 = E(XY)$$

وبذلك نستطيع حساب التغاير إلى X و Y كما يأتي:

$$cov(X,Y) = E(XY) - E(X)E(Y) = 3 - 1(2) = 1$$

ولحساب التباين لكل من X و Y فان:

$$var(X) = E(X^2) - \{E(X)\}^2 = 2 - \{1\}^2 = 1$$

$$var(Y) = E(Y^2) - \{E(Y)\}^2 = 6 - \{2\}^2 = 2$$

سنثبت نظرية مهمة جداً تربط دالة العزم المولدة المشتركة بمفهوم الاستقلالية للمتغيرات العشوائية.

نظرية 6-7

ليكن X و Y متغيرين عشوائيين لهما دالة التوزيع الاحتمالي المشتركة $f(x,y)$ ودالتي التوزيع الهامشية $f_1(x)$ و $f_2(y)$ على التوالي، ولتكن $M_{X,Y}(t_1,t_2)$ تمثل دالة العزم المولدة المشتركة للتوزيع. فان X و Y متغيرين عشوائيين مستقلين اذا وإذا فقط كان:

$$M_{X,Y}(t_1,t_2) = M_{X,Y}(t_1,0)\,M_{X,Y}(0,t_2) \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7\text{-}36)$$

البرهان

نفرض أن X و Y متغيرين عشوائيين مستقلين فان:

$$M_{X,Y}(t_1,t_2) = E(e^{t_1X+t_2Y}) = E(e^{t_1X})\,E(e^{t_2Y})$$

$$= M_{X,Y}(t_1,0)\,M_{X,Y}(0,t_2)$$

وهذا يعني انه اذا كان كل من X و Y متغيرين عشوائيين مستقلين فان دالة العزم المولدة المشتركة يمكن تحليلها إلى حاصل ضرب دالتي العزم المولدة الهامشية .

ولبرهان الاتجاه الآخر نفرض أن دالة العزم المولدة المشتركة تكتب بالشكل الآتي:

$$M_{X,Y}(t_1,t_2) = M_{X,Y}(t_1,0)\,M_{X,Y}(0,t_2)$$

في حالة كون X و Y متغيرين عشوائيين مستمرين فان كل واحد منهما له دالة عزم مولدة وحيدة وكما يأتي:

$$M_{X,Y}(0,t_2) = \int_{-\infty}^{\infty} e^{t_2 y} f_2(y)dy \quad \text{و} \quad M_{X,Y}(t_1,0) = \int_{-\infty}^{\infty} e^{t_1 x} f_1(x)dx$$

وبالتعويض عن هاتين الدالتين في دالة العزم المولدة المشتركة نحصل على:

$$M_{X,Y}(t_1,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1 x + t_2 y} f_1(x) f_2(y)\,dxdy$$

وبما أن $M_{X,Y}(t_1,t_2)$ هي دالة العزم المولدة المشتركة فإنها حسب التعريف تكون:

$$M_{X,Y}(t_1,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1 x + t_2 y} f(x,y)\,dxdy \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(7\text{-}37)$$

وعلية من وحدانية دالة العزم المولدة المشتركة نحصل على:

$$f(x,y) = f_1(x) f_2(y) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7\text{-}38)$$

لجميع قيم x و y .

وهذا يؤدي إلى أن X و Y متغيرين عشوائيين مستقلين.

مثال 7-16

ليكن X و Y متغيرين عشوائيين متقطعين لهما دالة التوزيع الاحتمالي المشتركة الآتية:

$x \backslash y$	-1	0	1	2
-2	0	$\frac{1}{6}$	0	0
-1	$\frac{1}{3}$	0	0	$\frac{1}{6}$
0	$\frac{1}{6}$	0	$\frac{1}{6}$	0

إحسب ما يأتي:

- دالة العزم المولدة المشتركة إلى X و Y؟
- دالتي العزم المولدة الهامشية؟
- التغاير بالنسبة إلى X و Y؟

الحل

نحسب دالة العزم المولدة المشتركة وكمايلي:

$$M_{X,Y}(t_1,t_2)=\frac{1}{6}e^{-2t_1}+\frac{1}{3}e^{-(t_1+t_2)}+\frac{1}{6}e^{2t_2-t_1}+\frac{1}{6}e^{-t_2}+\frac{1}{6}e^{t_2}$$

ولحساب دالتي العزم الهامشية:

$$M_{X,Y}(t_1,0)=\frac{1}{3}+\frac{1}{2}e^{-t_1}+\frac{1}{6}e^{-2t_1}$$

$$M_{X,Y}(t_1,0)=\frac{1}{6}+\frac{1}{2}e^{-t_2}+\frac{1}{6}e^{t_2}+\frac{1}{6}e^{2t_2}$$

ولحساب التغاير فان:

$$\frac{\partial M_{X,Y}(t_1,t_2)}{\partial t_1}=-\frac{1}{3}e^{-2t_1}-\frac{1}{3}e^{-(t_1+t_2)}-\frac{1}{6}e^{2t_2-t_1}$$

$$\frac{\partial M_{X,Y}(t_1,t_2)}{\partial t_2}=-\frac{1}{3}e^{-(t_1+t_2)}+\frac{1}{3}e^{2t_2-t_1}-\frac{1}{6}e^{-t_2}+\frac{1}{6}e^{t_2}$$

$$\frac{\partial^2 M_{X,Y}(t_1,t_2)}{\partial t_1\partial t_2}=\frac{1}{3}e^{-(t_1+t_2)}-\frac{1}{3}e^{2t_2-t_1}$$

وبوضع $t_1 = t_2 = 0$ نحصل على :

$$\mu'_{1,0} = \left[\frac{\partial M_{X,Y}(t_1,t_2)}{\partial t_1}\right]_{t_1=t_2=0} = -\frac{5}{6}$$

$$\mu'_{0,1} = \left[\frac{\partial M_{X,Y}(t_1,t_2)}{\partial t_2}\right]_{t_1=t_2=0} = 0$$

$$\mu'_{1,1} = \left[\frac{\partial^2 M_{X,Y}(t_1,t_2)}{\partial t_1 \partial t_2}\right]_{t_1=t_2=0} = 0$$

وعليه فان التغاير يكون:

$$cov(X,Y) = \mu'_{1,1} - \mu'_{1,0}\mu'_{0,1} = 0$$

وبما ان $M_{X,Y}(t_1,t_2) \neq M_{X,Y}(t_1,0)M_{X,Y}(0,t_2)$ فان X و Y غير مستقلين.

ملاحظة:

من المثال السابق نلاحظ بان التغاير يساوي صفر وبالرغم من ذلك نلاحظ أن المتغيرين العشوائيين X و Y ليس مستقلين،مما يدل على انه إذا كان المتغيرين العشوائيين مستقلين فان التغاير لهما يساوي صفراً ولكن العكس ليس صحيح.

8-7 معامل الارتباط The Correlation Coefficient

سنثبت في هذه الفقرة مقياس مهم من المقاييس التي توضح العلاقة بين المتغيرات العشوائية التي سبق وان درسنا بعض تفاصيلها،وهذه العلاقة تسمى (معامل الارتباط)وهي توضح لنا مدى العلاقة الخطية بين المتغيرين العشوائيين X و Y .ولذلك سوف نقدم تعريفا لهذا المعامل ومن ثم نشتق بعض الخواص المتعلقة به.

تعريف 7-12

معامل الارتباط بين المتغيرين العشوائيين X و Y يرمز له بالرمز $\rho_{X,Y}$ ويعرف بالصيغة التالية:[2]

$$\rho_{X,Y} = \frac{E\left[\{X - E(X)\}\{Y - E(Y)\}\right]}{\{var(X)\,var(Y)\}^{1/2}} = \frac{cov(X,Y)}{\{var(X)\,var(Y)\}^{1/2}} \quad ...(7-39)$$

حيث أن $cov(X,Y)$ و $\text{var}(X)$ و $var(Y)$ موجودة وان $var(X) > 0$ و $var(Y) > 0$.

ويسمى أحيانا بمعامل ارتباط بيرسون(Pearson) ولتبسيط التعامل معه نرمز له بالرمز ρ ،ولدراسة هذا المعامل ومعرفة حقيقة مايقيسه سوف نتطرق إلى بعض الخواص المهمة له من خلال النظريات الآتية.

نظرية 7-7

اذا كان X و Y متغيرين عشوائيين مستقلين ، فان $\rho = 0$ ويقال عن المتغيرين بأنهما غير مرتبطين.

البرهان

من العلاقة $(7-39)$ والنظرية $(7-5)$ وبما أن X و Y متغيرين عشوائيين مستقلين فان $cov(X,Y) = 0$ وعلية فان $\rho = 0$.

ولابد من التأكيد بان العكس لهذه النظرية ليس صحيحاً بشكل عام،وذلك لأننا يمكن أن نحصل على $\rho = 0$ ولكن المتغيرين العشوائيين ليس مستقلين.(المثال 7-16)يوضح ذلك.

نظرية 7-8

لأي متغيرين عشوائيين X و Y فان معامل الارتباط ، اذا وجد، فتكون له القيمة بين -1 و 1 وبعبارة أخرى يكون $-1 \le \rho \le 1$.

البرهان

نفرض أن:

$$U = X - E(X)$$
$$V = Y - E(Y)$$

وان $\rho = \dfrac{E(UV)}{\{E(U^2)\,E(V^2)\}^{1/2}}$.

ولتكن الدالة التربيعية الآتية بالنسبة للمتغير t هي:

$$z(t)=E\left\{(U+tV)^{2}\right\}=E(U^{2})+2tE(UV)+t^{2}E(V^{2})\ ......(7\text{-}40)$$

من الواضح بان $z(t)\geq 0$ لجميع قيم t .

وإذا اعتبرنا الدالة التربيعية $z(t)=at^2+bt+c\geq 0$ لجميع قيم t. اذا كان $z(t)\geq 0$ لجميع قيم t، اذا كان المقدار $b^2-4ac\leq 0$ فانه لا يعطي حل حقيقي، وإذا لا فنحصل على جذرين حقيقيين للمعادلة $z(t)=0$.

وفي حالة $z(t)<0$ لبعض قيم t. نضع:

$$a=E(V^{2}),b=2E(UV),c=E(U^{2})$$

فنحصل على:

$$\{2E(UV)\}^{2}-4E(V^{2})E(U^{2})\leq 0.$$

أي أن:

$$4\{E(UV)\}^{2}\leq 4E(V^{2})E(U^{2})$$

وبالقسمة على $E(V^{2})E(U^{2})$ نحصل على:

$$\frac{\{E(UV)\}^{2}}{E(V^{2})E(U^{2})}=\frac{\{E[X-E(X)][Y-E(Y)]\}^{2}}{E(V^{2})E(U^{2})}\leq 1$$

وبأخذ الجذر ألتربيعي للبسط والمقام نحصل على:

$$\frac{\{E(UV)\}}{\{\ E(V^{2})E(U^{2})\ \}^{1/2}}=\frac{\{E[X-E(X)][Y-E(Y)]\}}{\{\ E(V^{2})E(U^{2})\ \}^{1/2}}=\rho^{2}\leq 1$$

ومنها نحصل على: $-1\leq\rho\leq 1$.

وهذا يبين بان قيم معامل الارتباط تقع بين 1-و1، ولكن المهم لنا هو ماذا يعني وقوع القيم ضمن هذا المدى ويجب ان نحصل على تفسير لذلك.

نظرية 7-9

اذا كان X و Y متغيرين عشوائيين وان $Y = \alpha + \beta X$ حيث ان α و β قيم حقيقية و $(\beta \neq 0)$، فان:

$\rho = 1$ اذا كان $\beta > 0$ و $\rho = -1$ اذا كان $\beta < 0$.

البرهان

بما أن $Y = \alpha + \beta X$ فان $E(Y) = \alpha + \beta E(X)$

وان $var(Y) = \beta^2 var(X)$.

كذلك فان :

وبما أن:

$$E(XY) = E\{X(\alpha + \beta X)\} = \alpha E(X) + \beta E(X^2)$$

$$\rho = \frac{E(XY) - E(X)E(Y)}{\{var(X)var(Y)\}^{1/2}}$$

$$= \frac{\alpha E(X) + \beta E(X^2) - E(X)\{\alpha + \beta E(X)\}}{\{var(X)\beta^2 var(X)\}^{1/2}}$$

$$= \frac{\beta\left[E(X^2) - \{E(X)\}^2\right]}{\left[\beta^2\{var(X)\}^2\right]^{1/2}} = \frac{\beta}{|\beta|} \qquad \text{(7-41)}$$

وبذلك نكون قد أثبتنا بان $\rho = 1$ اذا كان $\beta > 0$ و $\rho = -1$ اذا كان $\beta < 0$.

إن النظرية أعلاه تثبت لنا بأنه اذا كان للمتغيرين العشوائيين علاقة خطية تامة فان قيمة معامل الارتباط تكون ± 1، وان العكس لها هو أيضا صحيح ، أي أن اذا كان $\rho = \pm 1$ فان للمتغيرين علاقة خطية يرتبطان بها، أي أن $Y = \alpha + \beta X$.

كما نلاحظ بأنه اذا كان المتغيرين العشوائيين مستقلين فان معامل الارتباط يساوي صفراً ،وهذا يمثل متوسط القيمة المحسوبة لقيم النهايات التي يأخذها معامل الارتباط.

مثال 7-17

ليكن X و Y متغيرين عشوائيين مستمرين لهما دالة التوزيع الاحتمالي المشتركة الآتية:

$$f(x,y)=\begin{cases} e^{-(x+y)}\left\{1+\theta(2e^{-x}-1)(2e^{-y}-1)\right\}, & x,y\geq 0, |\theta|<1 \\ 0 & ,otherwise \end{cases}$$

أحسب مايلي:

- دالة التوزيع الاحتمالي الهامشية إلى X و Y ؟
- المتوسط والتباين للمتغيرين X و Y ؟
- معامل الارتباط بين المتغيرين X و Y ؟

الحل

لحساب الدالة الهامشية بالنسبة إلى المتغير X :

$$f_1(x)=\int_0^{\infty} f(x,y)\,dy=\int_0^{\infty}\left\{e^{-(x+y)}\left[1+\theta(2e^{-x}-1)(2e^{-y}-1)\right]\right\}dy$$

$$=e^{-x}\left\{\int_0^{\infty}e^{-y}dy+\theta(2e^{-x}-1)\int_0^{\infty}(2e^{-2y}-e^{-y})\,dy\right\}=e^{-x}, x>0$$

وهذا يمثل التوزيع الأسي القياسي حيث أن:

$$E(X)=var(X)=1$$

وبنفس الأسلوب نحصل على دالة التوزيع الهامشية بالنسبة للمتغير Y وهي:

$$f_2(y)=\int_0^{\infty} f(x,y)\,dx=\int_0^{\infty}\left\{e^{-(x+y)}\left[1+\theta(2e^{-x}-1)(2e^{-y}-1)\right]\right\}dx$$

$$=e^{-y}\left\{\int_0^{\infty}e^{-x}dx+\theta(2e^{-y}-1)\int_0^{\infty}(2e^{-2x}-e^{-x})\,dx\right\}=e^{-y}, y>0$$

ومنه نحصل على:

$$E(Y) = var(Y) = 1$$

وكذلك فان:

$$E(XY) = \int_0^\infty \int_0^\infty xyf(x,y)dxdy$$

$$= \int_0^\infty e^{-x} x dx \int_0^\infty e^{-y} y dy + \theta \int_0^\infty x(2e^{-2x} - e^{-x})dx \int_0^\infty y(2e^{-2y} - e^{-y})dy$$

$$E(XY) = (1)(1) + \theta(-\frac{1}{2})(-\frac{1}{2}) = 1 + \frac{1}{4}\theta$$

ومن العلاقة (39 – 7) وبما أن:

$$cov(X,Y) = E(XY) - E(X)E(Y) = (1 + \frac{1}{4}\theta) - (1)(1) = \frac{1}{4}\theta$$

وعليه فان معامل الارتباط هو:

$$\rho = \rho_{X,Y} = \frac{cov(X,Y)}{\{var(X)var(Y)\}^{1/2}} = \frac{\frac{1}{4}\theta}{1} = \frac{1}{4}\theta$$

وبما أن $|\theta < 1|$ فان $-\frac{1}{4} < \rho < \frac{1}{4}$.

ملاحظة:

في حالة التعامل مع معامل الارتباط كمعامل لعينة من المجتمع ، فاننا نفرض بان لدينا المشاهدات الآتية:

$(x_1, y_1), (x_2, y_2), \ldots\ldots (x_n, y_n)$ وهي تمثل مشاهدات مستقلة للمتغير العشوائي (X,Y).

اذا كان:

$$s_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \text{ و } \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \text{ و } \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

وان: $s_y{}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2$ و $s_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$

وعلية يمكن تعريف معامل الارتباط للعينة كما يأتي:

$$r = \frac{s_{xy}}{s_x s_y} \qquad \text{(7-42)}$$

حيث أن $s_x = \sqrt{s_x{}^2}$ و $s_y = \sqrt{s_y{}^2}$

حيث نسمي s_{xy} تغاير العينة(sample covariance) .

ومن خلال إجراء تعديل على تباين العينة باستبدال $(n-1)$ محل n يمكن لنا أن نضع صيغة سهلة للتعامل لحساب معامل الارتباط من خلال تعويض القيم أعلاه وإجراء بعض الاختصارات في العلاقة $(7-42)$ فنحصل على:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\;\bar{y}}{(\sum_{i=1}^{n} x_i{}^2 - n\bar{x}^2)^{1/2}(\sum_{i=1}^{n} y_i{}^2 - n\bar{y}^2)^{1/2}}$$

وبتعويض قيم $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ و $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ نحصل على:

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{\left\{ n\sum_{i=1}^{n} x_i{}^2 - (\sum_{i=1}^{n} x_i)^2 \right\}^{1/2} \left\{ n\sum_{i=1}^{n} y_i{}^2 - (\sum_{i=1}^{n} y_i)^2 \right\}^{1/2}} \qquad \text{(7-43)}$$

نلاحظ أن المعامل r تكون قيمه محصورة بين $1-$ و1، وهو بذلك يشابه إلى معامل الارتباط ρ .أي أن القيم ± 1 تحصل عندما تقع النقاط (x_i, y_i) على الخط المستقيم بشكل يجعلها تعبر عن هذا الخط بشكل كبير،وبمعنى آخر إنها تتقرب لتشكل هذا الخط وتكون العلاقة الخطية بين x و y .

مثال7-18

احسب معامل الارتباط لقيم المتغيرين العشوائيين X و Y الموضوعة في الجدول أدناه:

x	y	xy	x^2	y^2
9.8	7.6	74.48	96.04	57.76
10.4	8.6	89.44	108.16	73.96
10.2	7.9	80.58	104.04	62.41
8.4	7.6	63.84	70.56	57.76
11.7	8.6	100.62	136.89	73.96
9.7	8.4	81.48	94.09	70.56
9.6	7.6	72.96	92.16	57.76
9.3	7.4	68.82	86.49	54.76
10.6	8.7	92.22	112.36	75.69
9.7	7.1	68.87	94.09	50.41
10.6	8.7	92.22	112.36	75.69
9.8	7.8	76.44	96.04	60.85
12.6	9.0	113.40	158.76	81.00
132.4	105.0	1075.37	1362.04	852.56

وبتعويض المجاميع التي حصلنا عليها في كل عمود من الأعمدة أعلاه نحصل:

$$r = \frac{13(1075.37) - (132.4)(105.0)}{\{13(1362.04) - (132.4)^2\}^{1/2}\{13(852.56) - (105.0)^2\}^{1/2}}$$

$$= \frac{77.81}{(176.76)^{1/2}(58.28)^{1/2}} = +0.77$$

نلاحظ بان قيمة معامل الارتباط للعينة $r = +0.77$ تكون موجبة وهي تبين أن الارتباط بين المتغيرات العشوائية X و Y يكون ارتباطاً قوياً،وبما أن قيم معامل الارتباط القصوى تكون محصورة بين -1 و1،وان r لا يصل إلى هذه القيمة الا اذا كانت جميع النقاط تقع على الخط المستقيم وتمثله بشكل تام.

9-7 التوقع الشرطي Conditional Expectation

إن دراسة المتغيرات العشوائية ذات البعدين تجعلنا نهتم بدراسة التوقعات الشرطية لهذه المتغيرات ،حيث تحسب من خلال توزيعاتها الشرطية، فمثلاً دراسة العلاقات بين الزيادة في سكان مجتمع معين والارتفاعات الحاصلة فيها ومن ثم فهي تمثل التوقع أو متوسط هذه الزيادات في السكان.

تعريف 7-13

ليكن X و Y متغيرين عشوائيين ، ولتكن $g_1(x/y)$ تمثل دالة التوزيع الاحتمالي الشرطية(conditional p.d.f)بالنسبة للمتغير X عندما $Y = y$ معطى. فان قيمة التوقع الشرطي(conditional expected value) للمتغير X عندما $Y = y$ معطى هو:

$$E(X / Y = y) = \sum_{all\ x} x g_1(x / y) \qquad \text{(7-44)}$$

في حالة المتغير العشوائي من النوع المتقطع.

ويكون:

$$E(X / Y = y) = \int_{-\infty}^{\infty} x g_1(x / y) dx \qquad \text{(7-45)}$$

في حالة المتغير العشوائي من النوع المستمر.

وبالأسلوب نفسه يمكن لنا أن نعرف التوقع الشرطي للمتغير Y عندما $X = x$ معطى وهو:

$$E(Y / X = x) = \sum_{all\ y} y g_2(y / x) \qquad \text{(7-46)}$$

في حالة المتغير العشوائي المتقطع، أما في حالة كون المتغير العشوائي من النوع المستمر فان التوقع الشرطي للمتغير Y عندما $X = x$ معطى هو:

$$E(Y / X = x) = \int_{-\infty}^{\infty} y g_2(y / x) dy \qquad \text{(7-47)}$$

ولتعميم الحالات أعلاه،سنعتبر الدالة $h(X)$ هي دالة للمتغير X،وعليه نعرف التوقع الشرطي للدالة $h(X)$ اذا وجد بالصيغة الآتية:

$$E\{h(X)/y\} = \sum_{all\ x} h(x)g_1(x/y) \quad \text{.........................(7-48)}$$

في حالة كون المتغير العشوائي من النوع المتقطع ،أما في حالة المتغير العشوائي المستمر فيعرف التوقع بالصيغة الآتية:

$$E\{h(X)/y\} = \int_{-\infty}^{\infty} h(x)g_1(x/y)dx \quad \text{.........................(7-49)}$$

وبالأسلوب نفسه نستطيع حساب التوقع الشرطي للدالة $h(Y)$ كدالة للمتغير العشوائي Y .

وعندما تكون $h(X) = X^r$ فان التوقع الشرطي الذي يقابل هذه الدالة هو كالآتي:

$$E(X^r/Y=y) = \sum_{all\ x} h(x)g_1(x/y) \quad \text{.........................(7-50)}$$

ويسمى العزم الشرطي من الرتبة r للمتغير X حول نقطة الأصل عندما $Y = y$ معطى.

وبالأسلوب نفسه نحسب العزم الشرطي من الرتبة r للمتغير Y حول نقطة الأصل عندما $X = x$ معطى.

مثال 7-19

ليكن X و Y متغيرين عشوائيين متقطعين لهما دالة التوزيع الاحتمالي المشتركة الآتية(مثال 7-17):

$$f(x,y) = \begin{cases} e^{-(x+y)}\{1+\theta(2e^{-x}-1)(2e^{-y}-1)\}, & x,y \ge 0, |\theta| < 1 \\ 0 & ,otherwise \end{cases}$$

إحسب العزم الشرطي للمتغير X عندما $Y = y$ ؟

الحل

للحصول على دالة التوزيع الاحتمالي الشرطية للمتغير العشوائي X فان:

$$g_1(x/y)=\frac{f(x,y)}{f_2(y)}=\frac{e^{-(x+y)}\left\{1+\theta(2e^{-x}-1)(2e^{-y}-1)\right\}}{e^{-y}}$$
$$=e^{-x}\left\{1+\theta(2e^{-x}-1)(2e^{-y}-1)\right\}$$

وبذلك يكون العزم الشرطي من الرتبة r هو:

$$E(X^r/Y=y)=\int_0^{\infty}x^r g_1(x/y)dx$$
$$=\left\{1-\theta(2e^{-y}-1)\right\}\int_0^{\infty}x^r e^{-x}dx+2\theta(2e^{y}-1)\int_0^{\infty}x^r e^{-2x}dx$$
$$=r!\left\{1-\theta-(2e^{-y}-1)(1-2^{-r})\right\}$$

وبشكل خاص عندما يكون $r=1$ فان:

$$E(X/Y=y)=1+\frac{1}{2}\theta-\theta e^{-y}$$

ولحساب التباين الشرطي إلى X عندما Y معطى فان:

$$var(X/Y=y)=E(X^2/Y=y)-\left\{E(X/Y=y)\right\}^2$$
$$=1+\frac{1}{2}\theta-\frac{1}{4}\theta^2-(\theta-\theta^2)e^{-y}-\theta^2e^{-2y}$$

ملاحظة:

إن $E(X/Y=y)$ يمثل التوقع إلى X عندما يكون الحدث Y معطى،وهو بشكل عام دالة بالنسبة إلى y.

وتمثيله بيانياً يسمى انحدار المنحني(regression curve) من X على Y، وفي حالة $E(Y/X=x)$ فيسمى انحدار المنحني Y على X. والرسم البياني(7-7) أدناه يظهر هذه الحالة:

ملاحظة:

- نلاحظ من النظرية(7-10)بأنه اذا كان الانحدار خطياً و $\rho = 0$ فان قيمتي $E(X / y)$ و $E(Y / x)$ مستقلتين وتعتمدان على y و x على التوالي.
- بما أن $E\{h(X) / y\}$ يعتمد على y،حيث أن Y متغير عشوائي للمشاهدات،فان $E\{h(X) / y\}$ يمكن اعتباره نفسه متغير عشوائي ومن ثم يكون التوقع له هو $E\left[E\{h(X) / y\}\right]$، ويسمى التوقع المتكرر(Iterated Expectation).

نظرية 7-11

ليكن (X,Y) متغير عشوائي ذو بعدين و $h_1(X)$ و $h_2(Y)$ دوال للمتغيرين X و Y على التوالي. فان:

$$E\left[E\{h_1(X) / x\}\right] = E\{h_1(X)\} \quad \text{.............................(7-58)}$$

و

$$E\left[E\{h_2(Y) / y\}\right] = E\{h_2(Y)\}$$

البرهان

سوف نبرهن الجزء الأول معتبرين أن X و Y متغيرين عشوائيين مستمرين.

من التعريف فان:

$$E\{h_1(X) / y\} = \int_{-\infty}^{\infty} h_1(x) g_1(x / y) dx = \int_{-\infty}^{\infty} h_1(x) \frac{f(x, y) dx}{f_2(y)}$$

وبأخذ التوقع بالنسبة إلى Y، فان:

$$E\left[\ E\{h_1(X) / y\}\ \right] = \int_{-\infty}^{\infty} f_2(y) \left\{ \int_{-\infty}^{\infty} h_1(x) \frac{f(x, y) dx}{f_2(y)} \right\} dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_1(x) f(x, y) dx dy$$

ومن اتجاه آخر للتكامل فان:

$$E[\ E\{h_1(X)/y\}\ ]=\int_{-\infty}^{\infty}h_1(x)\left\{\int_{-\infty}^{\infty}f(x,y)dy\right\}dx$$

$$=\int_{-\infty}^{\infty}h_1(x)f_1(x)dx \quad \text{.....................(7-59)}$$

وبما أن $f_1(x)$ هي دالة التوزيع الاحتمالي الهامشية بالنسبة إلى X ، فان :

$$E\left[E\{h_1(X)/y\}\right]=\int_{-\infty}^{\infty}h_1(x)f_1(x)dx=E\{h_1(X)\}$$

وبذلك يكون البرهان قد تحقق.

مثال 7-20

اذا كان X و N متغيرين عشوائيين ، وان دالة التوزيع الاحتمالي للمتغير N هي $f(n)=p$ حيث أن $(N=n)$ وان المتوسط لها هو μ_N والتباين $\sigma_N{}^2$ و X له التوزيع الثنائي الشرطي $g(x/n)=\binom{n}{x}p^x(1-p)^{n-x}$ وان $E(X/n)=np$ وان $\mathrm{var}(X/n)=np(1-p)$ ، احسب المتوسط والتباين بالنسبة إلى X ؟

الحل

$$E(X)=E\{E(X/n)\}=\sum_{all\ n}f(n)E(X/n)$$

$$=\sum_{all\ n}f(n)np=p\mu_N$$

$$E(X^2)=E\{E(X^2/n)\}=\sum_{all\ n}f(n)E(X^2/n)$$

$$=\sum_{all\ n}f(n)\left[var(X/n)+\{E(X/n)\}^2\right]$$

$$=\sum_{all\ n}f(n)\{np(1-p)+n^2p^2\}$$

$$=p(1-p)\mu_N+p^2(\sigma_n{}^2+\mu_n{}^2)$$

$$var(X)=E(X^2)-\{E(X)\}^2=p(1-p)\mu_N+p^2\sigma_N{}^2$$

تمارين الفصل السابع

1. إذا كانت دالة التوزيع الاحتمالي المشتركة (j.p.d.f) للمتغيرين X و Y هي كما يأتي:

$x \backslash y$	-1	0	1	2
-1	0	$\frac{1}{6}$	$\frac{1}{6}$	0
0	$\frac{1}{6}$	0	$\frac{1}{6}$	0
1	0	0	0	$\frac{1}{6}$

اوجد: 1 - دالة العزم المولدة لكل من X و Y ؟

2- احسب $COV(X,Y)$ ؟

2. اثبت أن $\sigma^2 = npq$ في التوزيع الثنائي الحدين (Binomial dist.) ؟

3. إذا كانت دالة التوزيع الاحتمالي الثنائية $(Bivariate \quad p.d.f)$ للمتغيرين العشوائيين X و Y هي:

$$f(x) = \begin{cases} \frac{1}{9}x(y-x) & ,0 \le x \le 3 \quad 2 \le y \le 4 \\ 0 & ,otherwise \end{cases}$$

احسب : 1- $E(X+Y)$ ؟ 2- $E(X-Y)$ ؟ 3- احسب التباين للتوزيع الهامشي (Marginal Dist.) لـ X و Y ؟ 4- احسب معامل الارتباط (Correlation cof.) بين X و Y ؟

4. إذا كان X و Y متغيرين عشوائيين مستمرين $(c.r.v)$ لهما دالـة التوزيـع التجميعيـة الثنائية $(Bivariate \quad c.d.f)$ الآتية :

$$F(x,y)=\begin{cases} 0 & y,x \le 0 \quad ,y \le 0 \\ \dfrac{xy}{(1+x)(1+y)} & ,x,y>0 \end{cases}$$

اوجد 1- دالة التوزيع الاحتمالي المشتركة لـ x و y ؟

2- اوجد الدوال الهامشية (Marginal Fun.) لـ X و Y ؟

3- اثبت إن X و Y متغيرين عشوائيين مستقلين ؟

5. اذا كان المتغيرين العشوائيين X و Y لهما دالة التوزيع التجميعية المشتركة الآتية:

$$F(x,y)=\begin{cases} 0 & ,x \le 0 \quad or \quad y \le 0 \\ xy \big/ \{(1+x)(1+y)\} & ,x,y>0 \end{cases}$$

أوجد دالـة التوزيـع الاحتماليـة المـشتركة للمـتغيرين X و Y،ثـم أوجـد دوالهمـا التوزيعية الهامشية،وبين أن المتغيرين العشوائيين مستقلين؟

6. اذا كان المتغيرين العشوائيين X و Y لهما الدالة التالية:

$$f(x,y)=\begin{cases} x^3y+\dfrac{\lambda^3}{16}y^2 & ,0<x<1,0<y<2 \\ 0 & ,otherwise \end{cases}$$

تحقق من كون $f(x,y)$ هي دالـة توزيـع احتمـالي مـشتركة؟ ،ثـم اوجـد الـدوال التوزيعية الهامشية للمـتغيرين العـشوائيين ؟،ثـم اوجـد دالـة التوزيـع الاحتمـالي الشرطي إلى X عندما $Y=y$ معطى؟.

7. اذا كان X و Y متغيرين عشوائيين ،وان $U = \alpha_1 + \beta_1 X$ ، $V = \alpha_2 + \beta_2 Y$،حيث أن α_1 و α_2 و β_1 و β_2 ثوابت اختيارية،بين ان معامل الارتباط بين U و V يساوي معامل الارتباط بين X و Y ؟

8. اذا كان X و Y متغيرين عشوائيين مستمرين لهما دالة التوزيع الاحتمالية المشتركة التالية :

$$f(x,y)=\begin{cases} 1/(2x^2y) & ,1<x<\infty, 1/x<y<x \\ 0 & ,otherwise \end{cases}$$

اشتق دالة التوزيع الهامشية للمتغيرين X و Y ؟،ثم اوجد دالة التوزيع الاحتمالية الشرطية إلى X عندما $Y = y$ معطى؟