

الفصل الثالث

المتغير العشوائي والتوزيع الاحتمالي

(١-٣) مقدمة

رأينا أن التجربة هي أي عملية تؤدي إلى قياس أو ملاحظة. وعدد الخريجين من جامعة الملك سعود مثلا هو قياس كمي؛ أو ملاحظة كمية، وأن تفوز الفرس «روعة» في سباق نادي الفروسية القادم أولا تفوز ملاحظة وصفية أو كيفية. ويمكننا دائما رد المعلومات الكيفية إلى معلومات رقمية بتخصيص عدد لكل نتيجة وصفية وفق نظام متفق عليه سلفا، فנסجل، مثلا، الرقم 1 إذا ربحت «روعة» السباق والرقم 0 إذا لم تربحه. وإذا رمزنا لعدد الخريجين بـ X ، ولنتيجة «روعة» في السباق بـ Y ، فمع نهاية كل عام دراسي سنحصل على قيمة للمتغير X ، ومع ختام كل سباق تشارك فيه «روعة» سنحصل على قيمة لـ Y . ومن الطبيعي أن نقول عن متغير مثل X أو Y إنه متغير عشوائي، لأن القيم التي يفترضها كل منهما مرتبطة بتجارب عشوائية.

مثال (١-٣)

لتكن التجربة اختيارا عشوائيا لطالب من الطلاب المسجلين في جامعة الملك سعود، وليكن:

$$X = 1 \text{ أو } 0 \text{ وفقا لما إذا كان يسكن أو لا يسكن في المدينة الجامعية.}$$

$$Y = \text{عدد إخوته.}$$

$$Z = \text{طوله بالسنتيمتر.}$$

فالمغيرات X, Y, Z هي متغيرات عشوائية. ونلاحظ أن فضاء العينة لمثل هذه التجربة هو مجموعة الطلاب المسجلين في جامعة الملك سعود، كل طالب يمثل نقطة عينة (نتيجة ممكنة). وكل متغير من هذه المتغيرات الثلاثة يأخذ قيمة واحدة وواحدة فقط عند كل نقطة عينة: وهو من هذا الوجه يشكل دالة عددية معرفة على فضاء العينة. فمن أجل كل طالب يأخذ X قيمة واحدة فقط هي إما 1 أو 0، ويأخذ Y قيمة واحدة فقط هي عبارة عن عدد صحيح غير سالب. ويأخذ Z قيمة واحدة فقط هي عبارة عن عدد حقيقي موجب.

مثال (٣-٢)

لتكن التجربة هي قذف ثلاث قطع نقود، وليكن X عدد أوجه الـ H التي نحصل عليها. فالمتغير X هو متغير عشوائي قيمه الممكنة 0 أو 1 أو 2 أو 3. وهو يأخذ عند كل نقطة عينة من النقاط الثماني التي يتضمنها فضاء العينة لهذه التجربة قيمة واحدة فقط من هذه القيم الممكنة. والجدول (٣-١) يبين ذلك.

فضاء العينة S	(HHH)	(HHT)	(HTH)	(THH)	(HTT)	(THT)	(TTH)	(TTT)
X	3	2	2	2	1	1	1	0

ومن الواضح أن X يمثل دالة عددية معرفة على فضاء العينة S . $X \rightarrow \{0, 1, 2, 3\}$ وأن $X(HHH) = 3$ ، $X(HHT) = 2$ ، $X(HTH) = 2$ ، الخ... وما سبق يتضح لنا، بصورة عامة، التعريف التالي للمتغير العشوائي.

(٣-١-١) تعريف المتغير العشوائي

المتغير العشوائي هو دالة عددية معرفة على فضاء عينة. وقد رأينا في الفصل السابق أن الحادثة هي مجموعة جزئية من فضاء عينة، فما هو حكم $X = 2$ ، مثلاً، وهل تمثل حادثة؟ والجواب نعم لأن $X = 2$ تعني وقوع واحدة من النقاط (HHT) أو (HTH) أو (THH). وسنصطلح على أن $[X = 2]$ تمثل الصورة العكسية لـ $X = 2$ أو $X^{-1}(2)$ ونكتب:

$$[X = 2] = X^{-1}(2) = \{(HHT), (HTH), (THH)\}$$

وهذا يسمح لنا بالقول إن $X = 2$ حادثة عددية نعبر عنها بدلالة المتغير العشوائي . ذلك لأن لها ما يقابلها في فضاء العينة الأصلي S . ونلاحظ أكثر من ذلك أن الحوادث العددية $X = 0$ ، $X = 1$ ، $X = 2$ ، $X = 3$ هي حوادث متنافية وتشكل تجزئة لفضاء العينة الأصلي S ، ففي الواقع :

$X = 0$ تمثل الحادثة $\{(TTT)\} = X^{-1}(0)$ أي مجموعة نقاط العينة التي يأخذ فيها X القيمة 0 ،

$X = 1$ تمثل الحادثة $\{(HTT), (THT), (TTH)\} = X^{-1}(1)$ ، أي مجموعة نقاط العينة التي يأخذ فيها X القيمة 1 ،

$X = 2$ تمثل الحادثة $\{(THH), (HTH), (HHT)\} = X^{-1}(2)$ ، أي مجموعة نقاط العينة التي يأخذ فيها X القيمة 2 ،

$X = 3$ تمثل الحادثة $\{(HHH)\} = X^{-1}(3)$ ، أي مجموعة نقاط العينة التي يأخذ فيها X القيمة 3 .

وهي في مثالنا هنا ، وفي غيره أيضا ، متنافية بالضرورة ، لأنه لو كان بين أي اثنين منها نقطة عينة مشتركة ، لاقتضى ذلك أن يكون لـ X قيمتان مختلفتان في تلك النقطة ، مما يتناقض مع حقيقة أن X دالة كما ينص التعريف . وسنقول إن المتغير X ولد فضاء عينة جديدا هو مجموعة قيمه الممكنة $\{0, 1, 2, 3\}$.

مثال (٣-٣)

نقذف قطعة نقود حتى يظهر وجه الـ H للمرة الأولى . وليكن X عدد القذفات التي نحتاجها . النتائج الممكنة للتجربة أو فضاء العينة هو :

$$H, TH, TTH, TTT, \dots$$

ومن الواضح أن X يمكن أن يكون 1 أو 2 أو 3 الخ . . . أي أن فضاء العينة الذي ولده X ، أو مجموعة قيمه الممكنة هي مجموعة الأعداد الطبيعية . $\{1, 2, 3, 4, \dots\}$.

(٣-٢) تصنيف المتغيرات العشوائية

ولنعد إلى المثال (٣-١) ، ولنتساءل عن مجموعة القيم الممكنة لـ Z ، طول الطالب . بها أننا سنستخدم مسطرة مدرجة لقياس الطول فإن طول الطالب سيقابل

نقطة على هذه المسطرة هي، في الواقع، نقطة على محور موجه. والقيمة التي يأخذها Z يمكن أن تكون أي نقطة من فترة على محور موجه. وبالطبع يوجد في أي فترة من محور موجه، مهما كانت صغيرة، ما لا نهاية له ولا يمكن عدّه أو حصره من النقاط. وبالرغم من أن فضاء العينة الذي يولده المتغير X في المثال (٣-٣) لا نهائي أيضا. إلا أن هناك خلافا أساسيا بين طبيعتي الفضاءين. فنقاط فترة من محور الأعداد الحقيقية هي مجموعة لا نهائية لا يمكن عدّها، أي لا يمكن إقامة تقابل بين هذه النقاط وبين الأعداد الصحيحة الموجبة $1, 2, 3, \dots$. ولو أخذنا الفترة $[160, 200]$ ، مثلا، واعتبرنا 160 مقابلا للعدد الصحيح 1، ثم سألنا أنفسنا ما هو العدد الذي يليه أي العدد الذي سيقابل 2 لاستحالت الإجابة. ومهما كان العدد الذي نرشحه قريبا من 160 فسبقى بين مثل هذا العدد والـ 160 ما لا يحصى ولا يعد من الأعداد. أما قابلية العد في فضاء العينة المتولد عن المتغير X في المثال (٣-٣) فهي أمر واضح لا يحتاج إلى تعليق. وهكذا نجد أن قابلية العد تميز بين صنفين من فضاءات العينة سنعرفهما فيما يلي:

(٣-٢-١) الفضاء المنفصل

نقول عن فضاء عينة إنه فضاء منفصل إذا كان مجوي عددا متتهيا من النقاط أو لا نهاية قابلة للعد من النقاط.

(٣-٢-٢) الفضاء المتصل

نقول عن فضاء عينة إنه فضاء متصل (أو مستمر) إذا كان مجوي لا نهاية غير قابلة للعد من النقاط.

ووفقا لهذا التصنيف نصنف المتغيرات العشوائية إلى متغيرات منفصلة ومتغيرات متصلة (أو مستمرة).

(٣-٢-٣) المتغير العشوائي المنفصل

نقول إن المتغير العشوائي منفصل إذا كانت مجموعة قيمه الممكنة مجموعة متتهية أو لا نهائية قابلة للعد. أي إذا كان فضاء العينة الذي يولده هذا المتغير فضاء منفصلا.

(٣-٢-٤) المتغير العشوائي المتصل (المستمر)

نقول إن المتغير العشوائي متصل (أو مستمر) إذا كانت مجموعة قيمه الممكنة لا نهائية وغير قابلة للعد. أي إذا كان فضاء العينة الذي يولده هذا المتغير متصلا (أو مستمرا).

(٣-٣) المتغيرات العشوائية المنفصلة وتوزيعاتها الاحتمالية

رأينا أنه يمكن التعرف على نوع المتغير العشوائي من خلال اتصاف مجموعة قيمه الممكنة أو عدم اتصافها بقابلية العد. وإذا أمكن للمتغير أن يفترض أو يتخذ عددا متتهيا أو لا نهائيا قابلا للعد من القيم فهو متغير منفصل. وفي معظم المسائل التي نواجهها في الحياة العملية تمثل المتغيرات المنفصلة قياسات على شكل تعداد مثل عدد حوادث المرور في مدينة الرياض خلال أسبوع، وعدد الإشارات الحمر التي تواجهها في طريقك إلى عملك، وعدد القطع المعيبة صناعيا في الإنتاج اليومي لمصنع، وعدد حالات الطلاق خلال سنة في مدينة معينة، وعدد البكتريا في ستمتر مكعب من الماء، وعدد الطائرات التي تصل في اليوم في رحلات دولية إلى مطار الملك خالد الدولي إلخ. وإذا كان عدد الإشارات التي تجتازها في طريقك إلى عملك هو عشر إشارات فإن عدد الإشارات الحمر التي يمكن أن تواجهها يتراوح بين 0 و 10 وعدد البكتريا X في ستمتر مكعب من الماء يمكن أن يكون كبيرا جدا إلا أنه محدود على أي حال، أي أن $X = 0, 1, \dots, n$ حيث n عدد كبير جدا.

ودالة التوزيع لمتغير عشوائي منفصل هي صيغة أو جدول يعرض القيم الممكنة والاحتمال الموافق لكل قيمة.

مثال (٣-٤)

في المثال (٣-٢) أوجد التوزيع الاحتمالي لـ X .

الحل

بالعودة إلى فضاء العينة الأصلي للتجربة وهو الفضاء المذكور في الجدول (٤-١) ومن اتزان أو تناظر قطع النقود، يمكننا إقامة نموذج احتمالي على هذا

الفضاء بتوزيع حصص متساوية على النقاط الثماني التي يتضمنها فضاء العينة. وبذلك يكون الاحتمال الموافق لكل نقطة هو $1/8$. وبعد أن نبني نموذجا احتماليا على فضاء العينة الأصلي، يمكننا الإجابة عن احتمال أي حادثة في هذا الفضاء (أي مجموعة جزئية من هذا الفضاء). ولكننا هنا في صدد الإجابة عن حادثة عددية معبر عنها بدلالة المتغير العشوائي X . مثلا، ما احتمال أن يأخذ X القيمة واحد. ونكتب ذلك رمزيا $P(X = 1)$ ، لنصطلح على ما تمليه البداهة هنا. وهو أن هذا الاحتمال هو احتمال الحادثة في فضاء العينة الأصلي التي تمثلها عبارة $X = 1$ ، أي احتمال الحادثة $\{(THT), (TTH), (HTT)\}$ وهو كما نعلم مجموع احتمالات النقاط الثلاث التي تتضمنها هذه الحادثة. أي $1/8 + 1/8 + 1/8 = 3/8$. وباختصار أكثر نقول إن احتمال $X = 1$ هو مجموع احتمالات نقاط العينة التي افترض فيها X القيمة 1. وبتطبيق هذه القاعدة على بقية القيم الممكنة نجد الجدول (٣-٢) حيث $f(x) = P(X = x)$.

جدول (٣-٢)

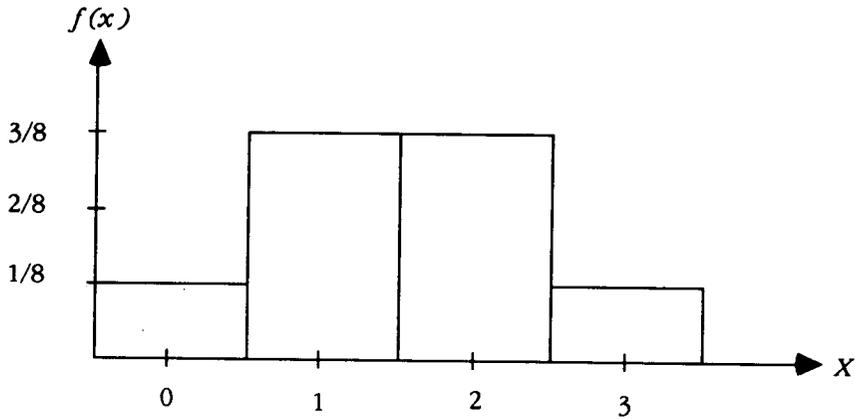
دالة التوزيع الاحتمالي لعدد أوجه الـ H عند قذف ثلاث قطع نقود.

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

ونلاحظ أولا أن مجموع الاحتمالات في هذا الجدول تساوي الواحد تماما. وهذه النتيجة متوقعة طالما أن الحوادث العددية التي تمثلها القيم المختلفة لـ X هي حوادث متنافية وتشكل تجزئة لفضاء العينة الأصلي S ، كما أوضحنا في المثال (٣-٢). وفي هذا المثال يمكننا تلخيص الجدول (٣-٢) بصيغة (علاقة) هي:

$$f(x) = \frac{\binom{3}{x}}{8}, \quad x = 0, 1, 2, 3$$

ويمكن تمثيل هذا التوزيع بيانيا لنحصل على ما يسمى بالمدرج الاحتمالي. فلتنخذ القيم الممكنة مراكز لفترات تمتد بمقدار الواحد (نصف على يمين القيمة ونصف على يسارها) ولنرسم فوق كل فترة مستطيلا ارتفاعه يساوي الاحتمال الموافق فنحصل على مدرج الاحتمال كما في الشكل (٣-١).



شكل (١-٣) المدرج الاحتمالي للتوزيع في المثال (٣-٤).

وبصورة عامة، عندما نبني نموذجاً احتمالياً على فضاء عينة S يمكننا استنتاج دالة التوزيع الاحتمالي لأي متغير عشوائي، X مثلاً، معرف على S . وذلك وفقاً للقاعدة التالية:

مجموع احتمالات نقاط العينة التي أخذ فيها X القيمة x $f(x) = P(X=x)$ والتوزيع الاحتمالي لمتغير عشوائي منفصل ليس إلا نموذجاً احتمالياً نقيمه على فضاء العينة الذي ولده هذا المتغير العشوائي. وإذا تذكرنا الشروط التي يجب أن يحققها نموذج احتمالي كما وردت في الفقرة (٢-٩) يمكن أن نستنتج هنا القاعدة التالية:

- يجب أن تحقق دالة التوزيع $f(x)$ لمتغير عشوائي منفصل X الشرطين التاليين:
- ١- $f(x) \geq 0$ مهما تكن x .
 - ٢- $\sum_x f(x) = 1$ حيث \sum_x تعني المجموع فوق جميع القيم الممكنة x .

مثال (٣-٥)

أوجد التوزيع الاحتمالي للمتغير X في المثال (٣-٣).

ومن هذا النموذج المعطى في الجدول (٣-٣) نستنتج بسهولة، وبتطبيق القاعدة العامة المعطاة أعلاه، دالة التوزيع الاحتمالي للمتغير X كما في الجدول (٣-٤). ويمكن التعبير عن هذه الدالة بالصيغة التالية:

جدول (٣-٣)

النموذج الاحتمالي على فضاء العينة الأصلي

نقطة العينة	الاحتمال الموافق
H	1/2
TH	1/4
TTH	1/8
TTHH	1/16
:	:
:	:
:	:

جدول (٤-٣)

دالة التوزيع الاحتمالي لـ X

X	f(x)
1	(1/2)
2	(1/2) ²
3	(1/2) ³
4	(1/2) ⁴
:	:
:	:
:	:

$$f(x) = \left(\frac{1}{2}\right)^x \cdot x = 1, 2, 3, \dots$$

وللتحقق من أن الدالة $f(x)$ تحقق شرطي دالة التوزيع المذكورين أعلاه ، نلاحظ أولاً أن جميع قيم الدالة غير سالبة وأن

$$\sum_x f(x) = \sum_{x=1}^{\infty} \left(\frac{1}{2}\right)^x = \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \dots$$

$$= \frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1.$$

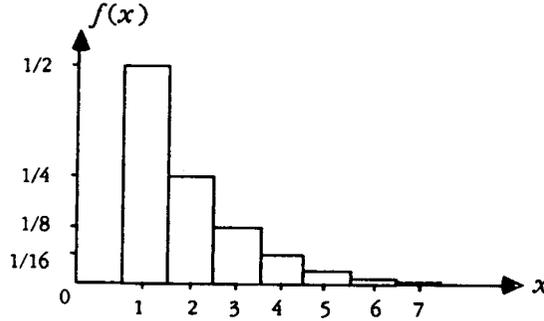
مثال (٦-٣)

ليكن X عدد النقاط على الوجه الظاهر عند رمي حجر نرد متماثل (متناظر). ما هي دالة التوزيع الاحتمالي لـ X؟

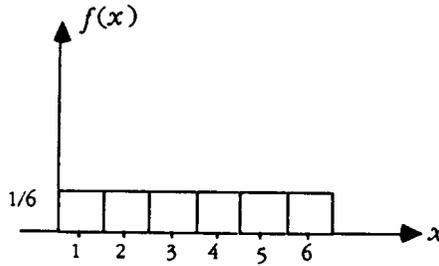
ويمكن التعبير عن الدالة هنا بالصيغة التالية:

$$f(x) = \frac{1}{6}, x = 1, 2, 3, 4, 5, 6.$$

ومن الواضح أن $f(x)$ تحقق شرطي دالة التوزيع . ويسمى مثل هذا التوزيع بالتوزيع المنتظم لأن الواحد موزع بالتساوي بانتظام) على كافة القيم الممكنة لـ X . والمدرج الاحتمالي مرسوم في الشكل (٣-٣) .



شكل (٣-٢) المدرج الاحتمالي للتوزيع في المثال (٣-٥)



شكل (٣-٣) المدرج الاحتمالي للتوزيع في المثال (٣-٦)

(٣-٤) التفسير العملي للتوزيع الاحتمالي لمتغير عشوائي منفصل

سنتصور مع كل متغير عشوائي ، X مثلاً، مجتمعا من القياسات ، هو المجتمع الناشئ عن تكرار قياس X عددا هائلا (لا نهائيا) من المرات . والتوزيع الاحتمالي لـ X يقدم وصفا لبنية هذا المجتمع . ففي المثال (٣-٤) يقول لك التوزيع الاحتمالي :

لو أنك كررت تجربة قذف ثلاث قطع نقود متماثلة (متناظرة) بلا حدود، وفي كل تكرار سجلت قيمة X (عدد أوجه الـ H التي حصلت عليها) فسيقع ذلك المجتمع من

القياسات الذي تحصل عليه في أربع فئات هي ، فئة الصفر، وفئة ال-1 ، وفئة ال-2 ، وفئة ال-3 . ولو قمت بتصنيف وكتابة جدول التكرار النسبي لهذا المجتمع من القياسات ، فستجد أن التكرار النسبي لكل من فئتي الصفر والواحد هو 12.5% ، وأن التكرار النسبي لكل من فئتي ال-2 وال-3 هو 37.5% . وبعبارة أخرى ، لو أنك رسمت مدرج التكرار النسبي لهذا المجتمع من القياسات فستحصل على الصورة نفسها التي يقدمها لك المدرج الاحتمالي لتوزيع X . وفي المثال (٣-٥) يقول لك التوزيع الإحتمالي أنك لو كررت التجربة بلا حدود وسجلت في كل مرة عدد القذفات التي احتجت إليها حتى ظهور وجه ال- H للمرة الأولى ، فستجد في 50% من هذه التكرارات أنك احتجت لقذفة واحدة ، وفي 25% من التكرارات لقذفتين ، وفي 12.5% من التكرارات لثلاث قذفات ، وفي 6.25% لأربع قذفات ، وهكذا وفي المثال (٣-٦) يقول لك التوزيع الاحتمالي أنك لو كررت تجربة رمي حجر نرد عددا لا نهائيا من المرات ، فستظهر الأوجه الستة بتكرارات نسبية متساوية ، وكل منها يساوي $16\frac{2}{3}\%$.

وبالطبع فإن تكرار أي تجربة عددا لا نهائيا من المرات هو مجرد افتراض نظري ، أي أن المجتمع من القياسات الموافق لمتغير عشوائي ليس إلا مجتمعا تصوريا قائما في الذهن فقط . وفي الحقيقة ليس التوزيع الاحتمالي إلا تجريدا ذهنيا لحالة فيزيائية واقعية ، أي أنه يشكل نموذجا رياضيا ، ويقدم وصفا لمجتمع نظري بلغة الواقع ، لغة الإحصاء الوصفي التي قدمناها في الفصل الأول من هذا الكتاب . ويكون المدرج الاحتمالي ، بهذا المعنى ، هو مدرج التكرار النظري لمجتمع القياسات .

ويبرز هنا سؤال جوهري . إذا كيف نتحقق من أن هذه التجريدات الذهنية تقدم محاكاة ناجحة لعالم الواقع؟

وللإجابة عن هذا السؤال يمكننا ، كما هو الحال في العلوم التجريبية ، اللجوء إلى التجربة والمشاهدة . وإذا كان توليد مجتمع لا نهائي غير ممكن عمليا إلا أنه يمكن تكرار التجربة عددا كبيرا من المرات ثم تصنيف القيم التي نحصل عليها للمتغير العشوائي ثم مقارنة صورة مدرج التكرار النسبي بصورة المدرج الاحتمالي . ولو قمنا بذلك لرأينا

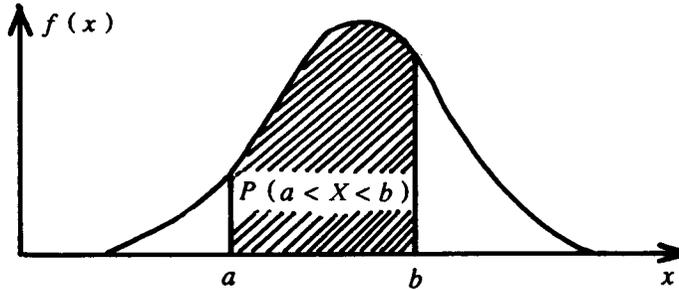
شبهها يثير الإعجاب ، حتى في حالة أعداد معتدلة لتكرارات التجربة . وهو شبه يزداد حدة ووضوحا مع زيادة عدد التكرارات . ويمكن للقارئ أن يقوم بتكرار أي من التجارب المذكورة في الأمثلة (٣ - ٤) ، (٣ - ٥) ، (٣ - ٦) ، مائة مرة ، مثلا ، ليحصل على مائة قياس للمتغير العشوائي X الذي نقيسه ، ويرسم لهذه القياسات المائة مدرج تكرار نسبي يقارنه بالمدرج الاحتمالي لـ X . وإذا كانت درجة التشابه بينهما لا ترضيه ، يمكنه زيادة عدد التكرارات مائة أخرى ورسم مدرج تكرار نسبي للمنتين من القياسات المتوفرة ، وسيلحظ أن الصورة الجديدة لمدرج التكرار النسبي قد اعتدلت في اتجاه المزيد من الشبه بين المشاهدة التجريبية والمقال النظري .

وعندما نقف بعد n من التكرارات ننظر إلى العدد المحدود من القياسات ، n ، على أنه عينة عشوائية من مجتمع القياسات الذي كنا سنحصل عليه لو استمر تكرار التجربة بلا حدود ، وهذه المقولة هي مقولة إصطلاحية في علم الإحصاء ولها فوائد جمة في التطبيقات العملية .

(٣ - ٥) المتغيرات العشوائية المتصلة

تشكل الكميات التي نستخدم للحصول على مقاديرها أجهزة قياس ، أو أدوات قياس ، متغيرات عشوائية متصلة . فالوزن والقوة والطول ومعدل هطول المطر ودرجة حرارة جسم ودرجة الامتحان لطالب كلها أمثلة على متغيرات عشوائية مستمرة . وقياسات مثل هذه المتغيرات هي نقاط على خط اتخذنا عليه تدريجا أو سلما للقياس ، أي أنها نقاط على المحور الموجه (محور الأعداد الحقيقية) ، أو على فترات من هذا المحور . ولا يمكننا ، في حالة متغير عشوائي مستمر ، تخصيص أي احتمال مهما كان صغيرا لأي قيمة من قيم المتغير نظرا للكثرة الكاثرة من القيم المختلفة ، إذ توجد لا نهاية غير قابلة للعد من النقاط في أي فترة مهما صغرت ، مما سيؤدي إلى الخروج على مسلمة الاحتمال الثانية . ولا بد من التفكير في طريقة لبناء النموذج الاحتمالي مختلفة تماما عما رأيناه في حالة متغير عشوائي منفصل .

لنعد بذاكرتنا الآن إلى مناقشة المضلع التكراري في الفقرة (١ - ٢ - ٣) ، حيث رأينا إمكانية تفسير المساحة تحت مدرج التكرار النسبي كاحتمال . وإلى منحني التكرار في الفقرة (١ - ٤) ، حيث تمثل كل نقطة على محور السينات قياسا ويمثل الإحداثي الصادي لتلك النقطة تواتر أو تكرار ظهور هذا القياس في المجتمع من القياسات الذي يصفه منحني التكرار . إذ تقدم لنا هذه الأفكار نقطة البداية في محاولة بناء نموذج احتمالي لمتغير عشوائي مستمر . لنبدأ بالقول إنه إذا كان تكرار ظهور القياس a ، مثلاً ، أكبر من تكرار ظهور القياس b ، فإن الكثافة الاحتمالية في a ينبغي لها أن تكون أكبر من الكثافة الاحتمالية في b . ولنعتبر منحني التكرار منحني كثافة يبين لنا كيف تتغير الكثافة الاحتمالية من نقطة إلى أخرى . ولنسمي الدالة المستمرة $f(x)$ ، التي بياناها هو منحني التكرار ، دالة كثافة احتمالية . وعندئذ ستمثل المساحات تحت هذا المنحني احتمالات . واحتمال أن يقع قياس المتغير x ضمن فترة (a, b) أي $P(a < X < b)$ هو المساحة تحت منحني الكثافة وفوق الفترة (a, b) . (انظر الشكل ٣ - ٤) .



شكل (٤ - ٣) دالة الكثافة الاحتمالية $f(x)$ أو منحني التكرار

أي أن $P(a < X < b)$ يساوي قيمة الدالة الأصلية لـ $f(x)$ محسوبة عند b مطروحا منها قيمة الدالة الأصلية عند a .

وترتب علينا مثل هذه الطريقة شرطين ، لا بد لأي دالة كثافة أن تحققهما ، كي لا نخرج على مسلمة الاحتمال . فما دام الاحتمال غير سالب ، لا يجوز أن يكون أي

جزء من منحنى الكثافة تحت المحور السيني . وبما أن احتمال الحادثة الأكيدة، أي $P(-\infty < X < +\infty)$ ، يجب أن يكون مساويا للواحد تماما، فإن المساحة تحت منحنى الكثافة بكامله يجب أن تساوي الواحد تماما . وهكذا نكتب القاعدة التالية :

(٣-٥-١) قاعدة

كي تصلح دالة متصلة $f(x)$ كدالة كثافة احتمالية يجب أن تحقق الشرطين التاليين :

$$١-٠ \geq f(x) \text{ لمهما يكن } x ،$$

٢- المساحة تحت بيان $f(x)$ (أي تحت منحنى الكثافة) تساوي الواحد تماما .

ووفقا لهذا التصور لو سألنا الآن عن احتمال أن يفترض متغير عشوائي متصل X قيمة محددة x ، مثلا لكان الجواب :

$$0 = \text{المساحة تحت منحنى الكثافة فوق النقطة } x = P(X = x)$$

والاحتمال صفر يعني الاستحالة . وهنا نجد أنفسنا في مأزق . لنفرض ، للتوضيح ، أن X يمثل طول إنسان ذكر بالغ بالستيمتر . فاستحالة أن يكون هذا المتغير مساويا لقيمة محددة ، أي قيمة ، تعني نفي وجود الجنس البشري ، وهي نتيجة في غاية السخف ، مما يثير الريبة في صلاحية النموذج الرياضي الموضوع لمتغير عشوائي متصل . ولكن لو تأملنا قليلا في هذه النتيجة لوجدنا أن التفسير المنطقي الوحيد لها هو أنه يستحيل على الإنسان أن يبتكر جهازا للقياس لا يخطئ ، أو أن يقيس بدون خطأ . ولا ريب أن لطول إنسان ذكر بالغ قيمة محددة تماما ، والمستحيل ليس وجود هذا القيمة وإنما القدرة على معرفتها ، أي أن ما يستحيل هنا هو الادعاء بأننا نستطيع قياس الأطوال بدون خطأ . ويجدر أن نقف قليلا أمام هذا المثال لنجد كيف يضطر المكابرون لتسجيل عجز الإنسان أمام باريه في شكل معادلة رياضية ، وفي ذلك آية لذوي البصيرة .

ويبقى سؤال وجيه آخر، إذ كيف نختار النموذج الموافق لحالة معينة؟ وما يمكن قوله هنا هو أن نستفيد من كل المعلومات المتوافرة لنا ثم نختار النموذج $f(x)$ وفق أفضل ما لدينا من قدرة على الحكم الصحيح . وتفرع المسألة هنا إلى مسألتين، فمثلا،

قد نعرف أن $f(x)$ على شكل جرس، ولكن من بين مثل هذه الأسرة من النماذج، ما هو على وجه التحديد، ذلك النموذج (ذلك المنحنى على شكل جرس) الذي يوافق الحالة المدروسة؟ وقد لا نعرف، على الوجه الآخر، حتى الشكل الأولي لـ $f(x)$ ، ونساءل، مثلاً، عما إذا كان ينبغي افتراضه على شكل جرس أم لا؟ ويتطرق الاستقراء الإحصائي إلى كل من المسألتين. ويقدم لنا الإحصاء الرياضي طرقاً لمعالجة مثل هذه المسائل سواء أكان المتغير X مستمراً أم منقطعاً. وبعد أن يقع اختيارنا على النموذج المناسب يمكننا في حالة متغير متصل حساب أي مساحة تحت منحنى الكثافة باستخدام الحساب التكاملي، وفي العديد من النماذج المعروفة والمستخدمه على نطاق واسع في طرق الإحصاء تتوافر جداول جاهزة تزودنا بمثل هذه المساحات.

ولكن هل يمكن الحصول على نتائج مفيدة باستخدام نماذج لم نتأكد تماماً من صحتها، أي من تمثيلها بصورة دقيقة للمجتمع المدروس؟ لننظر هنا إلى المهندس والكيميائي والفيزيائي وغيرهم، فنجد أن مختلف العلاقات العددية المستخدمة في مختلف فروع العلوم هي نماذج رياضية تقدم لنا تقريبات جيدة لواقع الحياة العملي. ويستخدم المهندس معادلاته لتحديد حجم وموضع دعامات جسر أو حجم وموضع جناح طائرة وما يهيمه، في المقام الأول، هو أن تقدم الجسور وأجنحة الطائرات الخدمات التي صممت من أجلها. وبالمثل فإن ما تقدمه الطرق الإحصائية من خدمات، هو المسطرة التي نقيس بها فائدة هذه الطرق، والقاعدة التي نحكم من خلالها على صحة ما تزودنا به من تنبؤات وقرارات تتعلق بالمجتمع المدروس. والجواب على سؤالنا نجده بوضوح في تطبيقات الإحصاء التي عم استخدامها وثبتت فوائدها، وقد اتسعت مساحة هذه التطبيقات لتشمل، على وجه التقريب، مختلف ميادين المعرفة ولتصبح بحق أداة رئيسة من أدوات الإنسان المعاصر في سعيه الدائم للكشف عن المجهول تحت شروط خاضعة للمصادفة.

(٦-٣) دالة التوزيع الاحتمالي المتجمع

رأينا في الفقرة (١-٣) أن التكرار المتجمع الصاعد يجيب عن السؤال التالي: ما هو التكرار النسبي لظهور قياس يقل عن قيمة محددة؟ وستجيب دالة التوزيع المتجمع

عن سؤال مشابه: ما احتمال أن يأخذ متغير عشوائي X قيمة أقل من أو تساوي قيمة محددة؟ وإذا رمزنا لهذه الدالة بـ F فإن قيمة F في نقطة x هي ببساطة احتمال أن يأخذ المتغير X قيمة أقل من أو تساوي x ، أي $P(X \leq x)$.

(٣-٦-١) حالة متغير عشوائي منفصل

دالة التوزيع المتجمع لمتغير عشوائي منفصل X ، دالة احتمالته $f(x)$ هي

بالتعريف:

$$F(t) = \sum_{x \leq t} f(x)$$

حيث $\sum_{x \leq t}$ تعني المجموع فوق جميع القيم الممكنة التي تقل عن t أو تساويها.

مثال (٣-٧)

في المثال (٣-٤) ما احتمال الحصول على وجه الـ H مرتين على الأكثر؟

الحل

المطلوب هو حساب

$$\begin{aligned} P(X \leq 2) &= F(2) = \sum_{x=0}^2 f(x) \\ &= f(0) + f(1) + f(2) \\ &= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8} \end{aligned}$$

تمرين

اعط تفسيراً عملياً لهذه النتيجة.

مثال (٣-٨)

في المثال (٣-٥) ما احتمال ألا يحتاج ظهور وجه الـ H للمرة الأولى إلى أكثر من

ثلاث قذفات؟

الحل

المطلوب هو حساب :

$$\begin{aligned} P(X \leq 3) &= F(3) = \sum_{x=1}^3 f(x) \\ &= f(1) + f(2) + f(3) \\ &= \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 = \frac{7}{8} \end{aligned}$$

تمرين

اعط تفسيراً عملياً لهذه النتيجة .

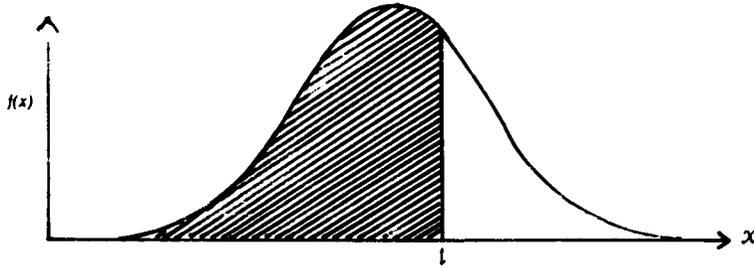
(٣-٦-٢) حالة متغير عشوائي متصل

دالة التوزيع المتجمع لمتغير عشوائي متصل X دالة كثافته الاحتمالية $f(x)$ ، هي

بالتعريف :

المساحة تحت منحنى الكثافة الواقعة إلى اليسار من النقطة t $P(X \leq t) = F(t) =$

انظر الشكل (٣-٥) .



شكل (٣-٥) : دالة التوزيع المتجمع لمتغير عشوائي متصل .

قلنا إن لكل متغير عشوائي مجتمعاً من القياسات هو المجتمع الناشئ عن تكرار قياس المتغير العشوائي مرة بعد أخرى إلى ما شاء الله . وأن التوزيع الاحتمالي يقدم وصفاً للبنية الداخلية لهذا المجتمع ويمثل التوزيع التكراري النظري له . وكما أن لكل مجموعة من القياسات مقاييس للترعة المركزية ومقاييس للتشتت فكذلك الأمر بالنسبة إلى مجتمع القياسات . وستحدث في الفقرتين القادمتين عن متوسط مجتمع القياسات وعن

تباينه، على الترتيب. وسنصطلح على استخدام عبارة «متوسط المجتمع» أو عبارة «متوسط التوزيع» لتعني الشيء نفسه. وكذلك سنقول في الوقت نفسه «تباين المجتمع» أو «تباين التوزيع»، و«الانحراف المعياري للمجتمع» أو «الانحراف المعياري للتوزيع». وسنرمز، كما جرت العادة في أدبيات الإحصاء، لمتوسط مجتمع بالحرف اليوناني μ (نطقه «ميو»)، وللانحراف المعياري لمجتمع بالحرف اليوناني σ (نطقه «سيجما»).

(٧-٣) التوقع الرياضي

(١-٧-٣) التوقع الرياضي لمتغير X

ليكن X متغيرا عشوائيا منفصلا دالة احتمالته $f(x)$. ولنرمز لتوقع X بـ $E(X)$ ،

فعندئذ:

$$E(X) = \sum_x x \cdot f(x)$$

حيث \sum_x يعني المجموع فوق كل القيم الممكنة للمتغير X .

هذا التعريف يقدم قاعدة لحساب توقع متغير عشوائي منفصل؛ إذ نضرب كل قيمة من القيم الممكنة للمتغير بالاحتمال المقابل لها ونجمع الجداءات الناتجة فنحصل على ما يسمى «بالقيمة المتوقعة رياضيا للمتغير»، أو اختصارا «القيمة المتوقعة للمتغير».

وكننا رأينا في الفقرة (١-٦-١) أن متوسط بيان مبوب يتضمن n قياسا هو:

$$\begin{aligned} \bar{y} &= \frac{\sum_{i=1}^m y_i f_i}{n} \\ &= y_1 \frac{f_1}{n} + y_2 \frac{f_2}{n} + \dots + y_m \frac{f_m}{n} \end{aligned}$$

حيث y_1, \dots, y_m هي القيم المختلفة التي يتضمنها البيان و f_1/n هو التكرار النسبي لـ y_1 و f_2/n هو التكرار النسبي لـ y_2 وهكذا. أي أنه لحساب المتوسط نضرب كل قيمة

بالتكرار النسبي لظهور هذه القيمة في البيان الإحصائي ثم نجمع الجداءات الناتجة . وبالعودة إلى التفسير العملي لدالة التوزيع الاحتمالي يتضح لنا أن $E(X)$ هو متوسط مجتمع القياسات . فنحن في عبارة $E(X)$ إنما نضرب كل قيمة من القيم المختلفة التي يتضمنها مجتمع القياسات بـ $f(x)$ الذي يمثل التكرار النسبي لظهور هذه القيمة في مجتمع القياسات . ويصبح المعنى التطبيقي لـ $E(X)$ أو التفسير العملي له واضحا . فالقيمة المتوقعة $E(X)$ للمتغير X هي متوسط القيم التي يمكن أن يفترضها هذا المتغير على المدى الطويل ، أي بعبارة أخرى متوسط مجتمع القياسات الموافق للمتغير X .

وتسمية $E(X)$ بالتوقع الرياضي ينبغي ألا تثير أي التباس إذ نستخدم في حسابه نموذجا رياضيا مجردا هو التوزيع الاحتمالي ، وهو يعبر ، في الحقيقة ، عن خاصية من خواص هذا النموذج الرياضي . إذ تمثل قيمة $E(X)$ الموضع أو النقطة على محور السينات (محو الأعداد) التي يتمركز حولها التوزيع الاحتمالي لـ X ، ولذلك سنسميها أيضا متوسط التوزيع الاحتمالي .

مثال (٣-٩)

في المثال (٣-٤) احسب $E(X)$.

الحل

$$E(X) = \sum_{x=0}^3 x \cdot f(x)$$

$$= 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{12}{8} = 1.5$$

ومما سبق يمكن القول إن :

(i) 1.5 هي القيمة المتوقعة رياضيا لعدد أوجه الـ H .

(ii) يتمركز التوزيع الاحتمالي لعدد أوجه الـ H حول النقطة 1.5 . ولو نظرنا إلى صورة المدرج الاحتمالي في الشكل (٤-١) لوجدنا أنه يتمركز حول النقطة 1.5 على محور الـ x . فالقيمة 1.5 هي متوسط التوزيع الاحتمالي .

(iii) التفسير العملي للقيمة 1.5 هو أنها تمثل القيمة المتوسطة لعدد أوجه الـ H على المدى الطويل (أي متوسط مجتمع القياسات). بمعنى أننا لو كررنا قذف قطع النقود الثلاث عددا هائلا من المرات وسجلنا في كل مرة عدد أوجه الـ H التي حصلنا عليها ثم حسبنا متوسط هذه الأعداد لحصلنا على 1.5 .

بعد أن تعلمنا كيفية حساب التوقع الرياضي لمتغير عشوائي X ، سندرس الآن طريقة حساب القيمة المتوقعة لدالة في X ، $g(X)$ مثلا . لنعد إلى المثال (٣ - ٤) ، ولنفرض أن $g(X) = X^2$. من الواضح أن X^2 يأخذ عند كل نقطة عينة في الجدول (٣ - ١) قيمة واحدة وواحدة فقط ، أي أنه دالة معرفة على فضاء عينة وبالتالي فهو متغير عشوائي . ويبين هذا ، بصورة عامة ، أن كل دالة في متغير عشوائي هي بدورها متغير عشوائي . ولكن كيف نحسب التوقع الرياضي لـ X^2 ؟ بما أن التوقع الرياضي يمثل متوسط مجتمع القياسات ، فلنبحث عن كيفية حساب متوسط مجتمع القياسات ومنها نستنتج قاعدة لحساب التوقع الرياضي . ولكن ما هو مجتمع القياسات الموافق لـ X^2 ؟ إنه بالضبط مجتمع القياسات لـ X بعد تربيع كل قيمة من قيمه . وإذا كان ظهور القيمة $X = 2$ يتكرر بنسبة $3/8$ ، كما نعلم من دالة التوزيع الاحتمالي لـ x ، فإن التكرار النسبي لظهور القيمة 4 في مجتمع القياسات الموافق لـ X^2 سيكون ، في مثالنا هنا ، $3/8$ أيضا ، وكذلك الأمر بالنسبة لبقية القيم . ولحساب القيمة المتوسطة ، على المدى الطويل ، نضرب كل قيمة ممكنة لـ X^2 بالتكرار النسبي لظهور هذه القيمة في مجتمع القياسات ثم نجمع النتائج ، أي :

$$\begin{aligned} \text{متوسط مجتمع القياسات لـ } X^2 &= 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8} \\ &= \frac{24}{8} = 3 \end{aligned}$$

وهذه النتيجة تسمح لنا بكتابة :

$$E(X^2) = \sum_{x=0}^3 x^2 \cdot f(x)$$

حيث $f(x)$ هي دالة التوزيع الاحتمالي للمتغير X .

وتقترح علينا المناقشة السابقة، بوضوح، التعريف التالي لتوقع دالة $g(X)$ ، بصورة عامة.

(٣-٧-٢) التوقع الرياضي لدالة عددية في X

ليكن X متغيراً عشوائياً منفصلاً، دالة توزيعه الاحتمالي $f(x)$ ، ولتكن $g(X)$ دالة عددية في x فعندئذ:

$$E[g(X)] = \sum_x g(x) \cdot f(x)$$

مثال (٣-١٠)

احسب $E(X^2)$ في المثال (٣-٦).

$$\begin{aligned} E(X^2) &= \sum_{x=1}^6 x^2 \cdot f(x) \\ &= 1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + 3^3 \times \frac{1}{6} + 4^2 \times \frac{1}{6} + 5^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6} \\ &= \frac{91}{6} = 15\frac{1}{6} \end{aligned}$$

وبصورة مشابهة تماماً نعرف توقع متغير عشوائي مستمر. كل ما في الأمر أن دالة الكثافة الاحتمالية تقوم مقام دالة التوزيع الاحتمالي للمتغير المنفصل وتصبح إشارة المجموع \sum إشارة تكامل \int وسوف لا نتطرق لذلك في هذا الكتاب.

ومن خواص إشارة المجموع \sum كما وردت في الفقرة (١-٥) يمكننا، بسهولة، التحقق من الخواص التالية لإشارة التوقع E .

(٣-٧-٣) خواص التوقع الرياضي

١- إذا كان c عدداً ثابتاً $[g(X)]$ في الفقرة (٣-٧-٢) تساوي مقدارا ثابتاً c

فإن:

$$E(c) = \sum_x c \cdot f(x) = c \sum_x f(x) = c$$

لماذا؟

٢- إذا كانت $g(X) = cX$ حيث c عدد ثابت فإن :

$$E(cX) = \sum_x cx f(x) = c \sum_x x f(x) = c \cdot E(X)$$

٣- إذا كان $g(X) = g_1(X) + g_2(X)$ فإن

$$\begin{aligned} E[g(X)] &= \sum_x g(x) \cdot f(x) \\ &= \sum_x [g_1(x) + g_2(x)] f(x) \\ &= \sum_x g_1(x) f(x) + \sum_x g_2(x) f(x) \quad (\text{استنادا إلى خواص } \sum) \\ &= E[g_1(X)] + E[g_2(X)] \end{aligned}$$

ومنه نستنتج الخاصة :

$$E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$$

وبصورة خاصة ، إذا كان X_1 و X_2 أي متغيرين عشوائيين فإن :

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

٤- من الخاصتين السابقتين يمكننا أن نكتب ، بصورة عامة ،

$$\begin{aligned} E[c_1X_1 + c_2X_2 + \dots + c_nX_n] \\ = c_1E(X_1) + c_2E(X_2) + \dots + c_nE(X_n) \end{aligned}$$

حيث X_1, X_2, \dots, X_n متغيرات عشوائية و c_1, c_2, \dots, c_n أعداد ثابتة .

مثال (٣-١١)

تقدم الإحصائية التالية وصفا لمجتمع الأسر التي تقطن مدنا كبيرة من حيث خاصية امتلاكها للسيارات :

20% من الأسر لا تمتلك أي سيارة و 50% من الأسر تمتلك سيارة واحدة و 15% من الأسر تمتلك سيارتين و 10% من الأسر تمتلك ثلاث سيارات و 5% من الأسر تمتلك

- أربع سيارات . إذا رمزنا بـ X لعدد السيارات وبـ Y لعدد العجلات التي تمتلكها أسرة .
 ١ - ما متوسط عدد السيارات للأسرة الواحدة؟
 ب - أحسب متوسط عدد العجلات للأسرة الواحدة .

الحل

- ١- من الواضح أن الوصف المعطى لمجتمع الأسر والمتعلق بقياسات X في هذا المجتمع يقدم دالة التوزيع الاحتمالي لـ X :

X	0	1	2	3	4
$f(x)$	0.20	0.50	0.15	0.10	0.05

ومتوسط عدد السيارات للأسرة الواحدة هو $E(X)$. ومن التعريف لدينا :

$$E(X) = 0(0.2) + 1(0.50) + 2(0.15) + 3(0.10) + 4(0.05) = 1.3$$

- ب - عدد العجلات عند أسرة هو عدد السيارات التي تمتلكها مضروباً بـ 5 . وإذا رمزنا لهذا المتغير العشوائي بـ Y فإن $Y = 5X$. والمطلوب هو $E(Y)$. ومن خواص التوقع لدينا :

$$E(Y) = E(5X) = 5E(X) = 5(1.3) = 6.5$$

ومتوسط عدد العجلات للأسرة الواحدة هو 6.5 عجلة .

- ذكرنا في مطلع هذه الفقرة أن لمجتمع القياسات الموافق لمتغير عشوائي X تبايناً وسنسمي مثل هذا التباين «تباين X » أو «تباين توزيع X » . ونذكر في الفقرة (١ - ٨ - ٤) أن تباين n من القياسات هو :

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

- ولأغراض تتعلق بالاستقراء الإحصائي نقسم على $(n-1)$ بدلاً من n عندما نحسب تباين عينة من القياسات). وعندما نكتب $\frac{1}{n} \sum_{i=1}^n \dots$ فهذا يعني أننا نجمع n مقادراً ثم نقسم على n ، أي نحسب متوسطاً . ويمكن التعبير عن التباين كلامياً كما يلي :

تباين n من القياسات هو متوسط مربعات انحرافات القياسات عن متوسطها .

وستبنى العبارة نفسها في تباين مجتمع من القياسات موافق لمتغير عشوائي X ، فمن المعروف أن متوسط هذا المجتمع هو $E(X)$ ، وأن مربع انحراف قياس عن المتوسط هو $[X - E(X)]^2$ ، والتباين ما هو إلا توقع هذا المقدار (أي قيمته المتوسطة) . ومنه التعريف التالي :

(٣-٧-٤) تباين متغير عشوائي

تباين متغير عشوائي X ، ونرمز له بـ $V(X)$ (أو σ_X^2) هو

$$V(X) = E[X - E(X)]^2$$

ومن خواص التوقع نستنتج الآن شكلا مختزلا أصح للحسابات .

وبغية الاختصار سنكتب μ بدلا من $E(X)$ فنجد :

$$V(X) = E(X - \mu)^2 = E(X^2 - 2\mu X + \mu^2)$$

$$= E(X^2) - 2\mu E(X) + E(\mu^2) ، \text{ (الخاصة الرابعة من خواص التوقع) ،}$$

$$= E(X^2) - 2\mu^2 + \mu^2 \text{ ، (الخاصة الأولى من خواص التوقع) ،}$$

$$= E(X^2) - \mu^2$$

ومنه الصيغة المختزلة للتباين :

$$V(X) = E(X^2) - [E(X)]^2$$

أو

$$V(X) = E(X^2) - \mu^2$$

أي أنه لحساب تباين X ، نحسب توقع مربع X ونطرح من الناتج مربع توقع X .

مثال (٣-١٢)

في المثال (٣-٤) ، احسب تباين X .

$$V(X) = E(X^2) - \mu^2$$

ولكن

$$E(X^2) = 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8} = \frac{24}{8} = 3$$

ونعلم من المثال (٣-٩) أن $\mu = E(X) = 1.5$ ، إذن :
 $V(X) = 3 - (1.5)^2 = 3 - 2.25 = 0.75$

(٣-٧-٥) الانحراف المعياري لمتغير

الانحراف المعياري لمتغير عشوائي X وسنرمز له بـ σ_x ، أو اختصارا σ عندما نأمن الالتباس ، هو الجذر التربيعي الموجب للتباين .

$$\sigma_x = \sqrt{V(X)}$$

وفي المثال السابق ، الانحراف المعياري لعدد أوجه الـ H الناتجة عن قذف ثلاث قطع متزنة من النقود هو

$$\sigma = \sqrt{0.75} = 0.865$$

خواص التباين

١ - تباين العدد الثابت هو الصفر .

$$V(c) = E(c^2) - [E(c)]^2 = c^2 - c^2 = 0$$

٢ - $V(cX) = c^2 V(X)$ حيث X أي متغير عشوائي و c عدد ثابت .

$$\begin{aligned} V(cX) &= E(cX)^2 - [E(cX)]^2 \\ &= c^2 E(X^2) - [cE(X)]^2 \\ &= c^2 [E(X^2) - (E(X))^2] \\ &= c^2 V(X) \end{aligned}$$

ونستنتج من هذه الخاصية أن :

$$\sigma_{cX} = |c| \sigma_X$$

أي إذا ضربنا المتغير X بعدد ثابت c فإن الانحراف المعياري لـ X يضرب أيضا بالعدد $|c|$.

٣ - يمكن البرهان أنه إذا كان المتغيران العشوائيان X_1 و X_2 مستقلين فيما بينهما فإن

$$V(X_1 + X_2) = V(X_1) + V(X_2)$$

ويمكن تعميم هذه القاعدة إلى أكثر من متغيرين فنقول إنه إذا كانت المتغيرات X_1, X_2, \dots, X_n مستقلة فيما بينها فإن :

$$V(X_1 + X_2 + \dots, X_n) = V(X_1) + V(X_2) + \dots + V(X_n)$$

ملاحظة

ما ذكرناه في الفقرة (١ - ٨) عن متباينة تشيبيشيف كان استعارة مبسطة لنظرية رياضية تحمل هذا الاسم وتتعلق بالتوزيعات الاحتمالية .

متباينة تشيبيشيف: ليكن X متغيرا عشوائيا متوسطه μ وانحرافه المعياري σ وليكن k أي عدد موجب ، فعندئذ :

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

أي

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

أي أن احتمال أن يأخذ X قيمة مختلف عن المتوسط μ بأقل من k انحرافا معياريا هو على الأقل $1 - \frac{1}{k^2}$. وبلغة هندسية نقول في حالة متغير عشوائي منفصل إن $1 - \frac{1}{k^2}$ على الأقل من المساحة تحت مدرج الاحتمال واقع بين $\mu - k\sigma$ و $\mu + k\sigma$. وفي حالة متغير عشوائي مستمر نقول إن ما لا يقل عن $1 - \frac{1}{k^2}$ من المساحة تحت منحنى الكثافة الاحتمالية واقع بين $\mu - k\sigma$ و $\mu + k\sigma$.

مثال (٣-١٣)

في المثال (٣-١١) احسب تباين X وانحرافه المعياري ، وتباين Y وانحرافه المعياري .

$$V(X) = E(X^2) - [E(X)]^2$$

وقد حسبنا في المثال (٣-١١) توقع X فوجدناه 1.3 ، ولدينا :

$$E(X^2) = 0^2 \times 0.20 + 1^2 \times 0.50 + 2^2 \times 0.15 + 3^2 \times 0.1 + 4^2 \times 0.05 = 3$$

ومنه :

$$V(X) = 3 - (1.3)^2 = 1.31$$

$$\sigma_X = \sqrt{V(X)} = \sqrt{1.31} = 1.1446$$

ومن الخاصة الثانية من خواص التباين نجد :

$$V(Y) = V(5X) = 25V(X) = 25 \times 1.31 = 32.75$$

$$\sigma_Y = \sqrt{32.75} = 5.72$$

أو

$$\sigma_Y = 5\sigma_X = 5(1.1446) = 5.723$$

تمارين (٣-١)

(١) في كل مما يلي حدد ما إذا كانت الدالة f تصلح دالة توزيع احتمالي لمتغير عشوائي

مجموعة قيمه الممكنة هي $\{1, 2, 3, 4\}$:

أ- $f(1) = 0.26, f(2) = 0.26, f(3) = 0.26, f(4) = 0.26$

ب- $f(1) = 0.15, f(2) = 0.28, f(3) = 0.29, f(4) = 0.28$

ج- $f(1) = \frac{1}{9}, f(2) = \frac{2}{9}, f(3) = \frac{1}{3}, f(4) = \frac{1}{3}$

د- $f(1) = 0.33, f(2) = 0.37, f(3) = -0.03, f(4) = 0.33$

ه- $f(1) = \frac{1}{4}, f(2) = \frac{3}{8}, f(3) = \frac{3}{16}, f(4) = \frac{5}{32}$

(٢) حدد ما إذا كانت الدوال التالية تصلح دوال توزيع احتمالي وعلل إجابتك :

أ- $f(x) = \frac{1}{5}, x = 1, 2, 3, 4, 5;$

ب- $f(x) = \frac{x+1}{4}, x = 1, 2, 3, 4;$

ج- $f(x) = \frac{x^2}{30}, x = 1, 2, 3, 4;$

د- $f(x) = \frac{x-2}{5}, x = 1, 2, 3, 4.$

وارسم المدرج الاحتمالي لكل دالة توزيع تجدها.

(٣) حزمة من البطاريات تتضمن 6 بطاريات ، اثنتان منها فاسدتان . اخترنا عشوائيا عينة من ثلاث بطاريات . إذا رمزنا بـ X لعدد البطاريات الفاسدة في العينة . أكتب التوزيع الاحتمالي لـ X ، وارسم المدرج الاحتمالي .

(٤) يتضمن صندوق أربع قطع صالحة وقطعة فاسدة . فحصنا هذه القطع واحدة فأخرى . وليكن X رقم الاختبار الذي عثرنا فيه على القطعة الفاسدة . اكتب توزيع X .

(٥) في كل من التمرينين (٣) و (٤) . أحسب متوسط التوزيع وتباينه .

(٦) قذفنا حجري نرد؛ وليكن X عد النقاط الظاهرة على الحجر الأول ، و Y عدد النقاط الظاهرة على الحجر الثاني .

ا- اكتب التوزيع الاحتمالي لـ $T = X + Y$ واحسب $E(T)$ ، $V(T)$.

ب- اكتب التوزيع الاحتمالي لـ $W = XY$ ، واحسب $E(W)$ ، $V(W)$.

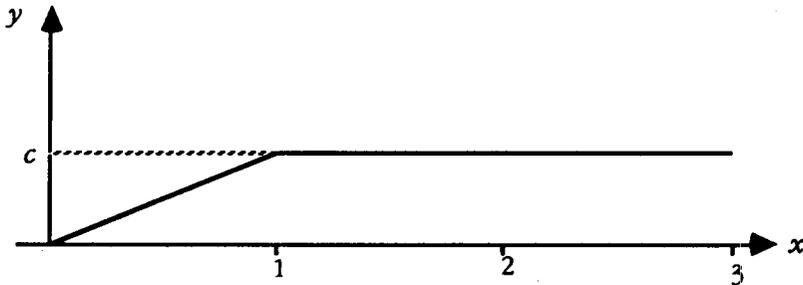
ج- مستخدما التوزيع الإحتمالي لكل من X و Y ، احسب $E(X)$ ، $E(Y)$ ،

$V(X)$ و $V(Y)$.

د- قارن بين $E(T)$ و $E(X) + E(Y)$ ؛ وبين $V(T)$ و $V(X) + V(Y)$.

(٧) في الشكل (٦-٣) المجاور، حدد قيمة c بحيث تصلح الدالة المرسومة في الشكل دالة كثافة احتمالية، واحسب:

$P(0.5 < X < 2.5)$ ، $P(X > 2.5)$ ، $P(X \leq 1.5)$ ، $P(X = 2)$ ، $P(X < 1.5)$



شكل (٦-٣)

- (٨) قذفنا ثلاث قطع نقود ، وليكن X عدد أوجه الـ H التي حصلنا عليها .
 ا- اكتب التوزيع الاحتمالي لـ X ، وارسم المدرج الاحتمالي .
 ب- احسب متوسط التوزيع $E(X)$ ، وتباين التوزيع $V(X)$.
 ج- نفذ هذه التجربة عمليا مائة مرة ، وسجل في كل مرة قيمة X ، ثم ارسم مدرج تكرار للقيم المائة لـ X . هل تجد أنه مشابه للمدرج الاحتمالي؟ استنتج من ذلك تفسيراً عملياً للمدرج الاحتمالي .
 د- احسب \bar{X} و S^2 متوسط وتباين القيم المائة لـ X التي حصلت عليها في ج- . هل يشكل \bar{X} تقديراً جيداً لـ $E(X)$ ، و S^2 تقديراً جيداً لـ $V(X)$ ؟

- (٩) مجتمع من خمسة أرقام هي $\{1, 2, 3, 4, 5\}$. سحبنا عشوائياً عينة من رقمين ، وليكن \bar{X} متوسط هذه العينة . اكتب التوزيع الاحتمالي لـ \bar{X} .

- (١٠) تاجر للمعدات الثقيلة يتصل في اليوم بزبون واحد أو زبونين ، وذلك باحتمال يساوي $1/3$ ، $2/3$ ، على الترتيب . وسيستج كل إتصال إما لا شيء ، أو صفقة بيع قيمتها خمسون ألف ريال ، وذلك باحتمال 0.9 ، 0.1 ، على الترتيب . احسب توقع مبيعاته اليومية .

- (١١) في طريقه إلى عمله ، يجتاز موظف ثلاث إشارات ضوئية . والإشارات تعمل مستقلة بعضها عن بعض . واحتمال أن يواجه إشارة حمراء هو 0.4 ، 0.8 ، 0.5 ، بالنسبة للإشارات الثلاث ، على الترتيب .

أ- ليكن Y عدد الإشارات الحمراء التي يواجهها الشخص في رحلته اليومية إلى عمله ، أوجد توزيع Y .

ب- احسب القيمة المتوقعة لـ Y وانحرافه المعياري .

ج- افترض أن وقت الانتظار لكل إشارة حمراء هو دقيقتان . ما هي القيمة المتوقعة والانحراف المعياري للوقت الذي ينتظره هذا الموظف للرحلة الواحدة .

- (١٢) لنقم بمحاكاة التجربة في التمرين الثالث بوضع علامات مميزة على ست قطع متماثلة من الورق ، بحيث تمثل اثنتان منها البطاريتين الفاسدتين ، وتمثل القطع

الأربع الباقية البطاريات الصالحة للاستعمال . ضع قطع الورق هذه في قبعة ،
أخلطها جيدا واسحب ثلاثا منها ، ثم سجل قيمة X عدد القطع التي تمثل
بطاريات فاسدة . أعد القطع إلى القبعة ثانية وكرر العملية نفسها من جديد حتى
تحصل على مائة قياس لـ X . ارسم مدرج التكرار النسبي لهذه العينة من القياسات
وقارنه مع المدرج الاحتمالي الذي حصلت عليه من ذلك التمرين .

(١٣) في التمرين الثالث أحسب $E(X) = \mu$ ، و $V(X) = \sigma^2$ ، وهما متوسط وتباين X في
المجتمع النظري من القياسات ، وذلك باستخدام دالة التوزيع التي حصلت عليها
هناك . ثم احسب المتوسط \bar{X} ، والتباين S^2 للعينة من القياسات التي حصلت
عليها في التمرين ١٢ ، هل يشكل \bar{X} تقديرا جيدا لـ μ ، و S^2 تقديرا جيدا لـ σ^2 ؟

(١٤) استخدم المدرج الاحتمالي الذي حصلت عليه في التمرين الثالث لحساب النسبة
من مجتمع القياسات الواقعة ضمن انحرافين معيارين على جانبي المتوسط ، وقارن
مع نظرية تشيبيشيف . أعد في عينة القياسات المذكورة في التمرين ١٢ .

(١٥) ولد عينة من 50 قياسا من المجتمع من القياسات الموافقة للمتغير X المذكور في
المثال (٦-٣) . وذلك بقذف حجر نرد 50 مرة ، وتسجيل X بعد كل قذفة .
احسب \bar{X} و S^2 للعينة ، وقارنها مع $E(X)$ و $V(X)$.