المحاضرة الثالثة

الإحصاء الحيوي

السنة الثالثة صيدلة

د. وديع علي

المتغيرات العشوائية وتوزيعاتها الاحتمالية

Random Variables and Probability Distributions

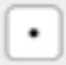
مقدمة:

تسمى نتائج التجربة العشوائية بالمتغيرات العشوائية.

المتغير العشوائي (Random Variable)

هو المقدار الذي يأخذ قيماً عددية مختلفة، والتي تعبر عن نتائج التجربة العشوائية .

فمثلاً عند إلقاء زهرة النرد مرة واحدة، فإن التجربة هنا عشوائية ونتائجها هي:

نتائج التجربة						
X	1	2	3	4	5	6

وتنقسم المتغيرات العشوائية إلى قسمين:

1) المتغير العشوائي المنفصل (المتقطع)
Discrete Random Variable

2) المتغير العشوائي المستمر (المتصل)
Continuous Random Variable

يقال إن المتغير العشوائي (X) متغير عشوائي منفصل إذا كان يأخذ قيماً تنتمي إلى مجموعة منتهية أو معدودة حتى ولو لم تكن هذه القيم صحيحة.

من الأمثلة على هذا المتغير :

- عدد الحوادث الشهرية على الطرق السريعة
- عدد الأسهم المخصصة للفرد في شركة مساهمة.

يقال إن المتغير العشوائي (X) متغير عشوائي متصل إذا كان يأخذ القيم الصحيحة والكسرية في مدى تغيره، أو إذا كان ينتمي إلى مجموعة غير منتهية أو معدودة.

من الأمثل على المتغير المتصل:

- أسعار المنتجات المختلفة
- أجور العمال في إحدى الشركات.

المتغيرات العشوائية والتوزيعات الاحتمالية

التوزيع الاحتمالي للمتغير العشوائي:

- عبارة عن دالة توضح احتمالات معينة لقيم المتحول العشوائي المختلفة.
- ويعبر عن هذه الدالة بجدول أو صيغة رياضية تبين قيم المتغيرات والاحتمالات المقابلة لكل منها.

فمثلا في تجربة إلقاء زهرة نرد نجد أن:

نتائج التجربة	⚀	⚁	⚂	⚃	⚄	⚅
X	1	2	3	4	5	6
$P(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

وتنقسم التوزيعات الاحتمالية إلى نوعين أساسيين هما:

- التوزيعات الاحتمالية المنفصلة
- التوزيعات الاحتمالية المتصلة.

1- التوزيعات الاحتمالية المنفصلة:

إذا كان X متحولاً عشوائياً منفصلاً يأخذ القيم $x_1, x_2, ..., x_n$ واحتمالات هذه القيم هي: $P(x_1), P(x_2), ..., P(x_n)$ فيقال إن المتغير العشوائي X يتبع توزيعاً احتمالياً منفصلاً $P(x)$ إذا تحققت الشروط الآتية:

- لجميع قيم X يكون $P(x) \geq 0$
- $\sum P(x) = 1$ (أي أن مجموع الاحتمالات يساوي الواحد).

خصائص أساسية للتوزيع الاحتمالي المنفصل:

- توقع التوزيع (التوقع الرياضي أو متوسط التوزيع):

$$E(X) = \mu = \sum xP(x)$$

- تباين التوزيع:

$$Var(X) = \sigma^2 = \sum x^2 P(x) - \mu^2$$

- الانحراف المعياري: $\sqrt{Var(X)} = \sigma = \sqrt{\sigma^2}$

مثال:

في شعبة لمادة الرياضيات تحوي 25 طالباً تم إجراء اختبار سريع من 5 درجات وكانت النتائج كما يلي:

3 طلاب حصلوا على 0

1 طالب حصل على 1

4 طلاب حصلوا على 2

8 طلاب حصلوا على 3

6 طلاب حصلوا على 4

3 طلاب حصلوا على 5

إذا كان المتغير العشوائي X يعبر عن علامة الطالب الذي يتم اختباره عشوائياً فأوجد:

1- التوزيع الاحتمالي لجميع قيم ذلك المتغير

2- المتوسط الحسابي والانحراف المعياري لهذا التوزيع الاحتمالي.

الحل:

1- إيجاد التوزيع الاحتمالي:

X	0	1	2	3	4	5
P(X)	0.12	0.04	0.16	0.32	0.24	0.12

2- إيجاد المتوسط الحسابي والانحراف المعياري:

X	P(X)	X.P(X)	X^2.P(x)
0	0.12	0	0
1	0.04	0.04	0.04
2	0.16	0.32	0.64
3	0.32	0.96	2.88
4	0.24	0.96	3.84
5	0.12	0.6	3
	Sum	2.88	10.4

$$\mu = \sum xP(x) = 2.88$$
$$\sigma^2 = \sum x^2 P(x) - \mu^2 = 10.4 - (2.88)^2$$
$$\sigma = \sqrt{10.4 - (2.88)^2} = 1.45$$

مثال:

إذا كان التوزيع الاحتمالي المنفصل لمتغير X هو كالتالي:

X	1	2	3	4
P(x)	k	0.2	0.6	0.1

المطلوب:

1- إيجاد قيمة k

2- حساب متوسط وتباين التوزيع.

الحل:

1- بما أن مجموع الاحتمالات يساوي الواحد فإن:

$$\sum P(x)=1;$$
$$k+0.2+0.6+0.1=1;$$
$$k=1-0.9=0.1$$

2- يمكن حساب متوسط وتباين التوزيع كما يلي:

X	1	2	3	4	Σ
$P(x)$	0.1	0.2	0.6	0.1	1
$x \cdot P(x)$	0.1	0.4	1.8	0.4	2.7
$x^2 \cdot P(x)$	0.1	0.8	5.4	1.6	7.9

$$\mu=E(X)=\sum xP(x)=0.1+0.4+1.8+0.4=2.7$$
$$\sigma^2=\sum x^2P(x)-\mu^2=7.9-(2.7)^2=7.9-7.29=0.61.$$

تمرين: اختر الإجابة الصحيحة:

1- أوجد القيمة المفقودة في الجدول الآتي:

x	0	1	2	3
P(x)	$\frac{1}{4}$	K	$\frac{1}{8}$	$\frac{1}{8}$

$A)\frac{1}{2}$ $B)\frac{1}{4}$ $C)\frac{1}{8}$ $D)\ 1.5$

2- المتوسط الحسابي للتوزيع السابق يساوي:

$A)\ 0.25$ $B)\ 0.125$ $C)\ 1.125$ $D)\ 1.5$

توزيع برنولي:

نسمي كل تجربة لها نتيجتان تجربة برنولي (ثنائية) نسمي إحدى النتيجتين نجاحاً والأخرى فشلاً. فإذا كان p يمثل احتمال النجاح فإننا نسمي p وسيط التجربة البرنولية و $q=1-p$ احتمال الفشل.

فعلى سبيل المثال إن إجراء عملية جراحية هي تجربة برنولية لها نتيجتان إما الشفاء (النجاح) أو تدهور الحالة الصحية (الفشل). ووسيط هذه التجربة p يتعلق بطبيعة العملية وحالة المريض.

تعريف:

يقال إن للمتحول العشوائي X توزيعاً برنولياً بالوسيط p إذا كان لـ X دالة الاحتمال:

$$f(x) = p^x q^{1-x} \; ; x = 0,1$$

أي أن لـ X جدول التوزيع الاحتمالي الآتي:

$X = x$	0	1
$f(x)$	q	p

حيث $p > 0$ و $q = 1 - p$.

بعض خواص توزيع برنولي:

1) $M_X(t) = E(e^{Xt}) = \sum_{k=0}^{\infty} f(x_k)e^{tx_k} = f(0)e^0 + f(1)e^t = q + pe^t$

2) $\varphi_X(t) = E\left(e^{iXt}\right) = q + pe^{it}$

3) $E(X) = \frac{dM_X(0)}{dt} = \left[pe^{it}\right]_{t=0} = p$

4) $\sigma_X^2 = \alpha_2 - \alpha_1^2 = p - p^2 = p(1-p) = pq$

5) $\sigma_X = \sqrt{pq}$

6) $\gamma = \frac{\mu_3}{\sigma^3} = \frac{q-p}{\sqrt{pq}}$

حيث:

$$\mu_3 = \alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^3$$

$$\alpha_3 = \frac{d^3}{dt^3}M_X(0) = [pe^t]_{t=0} = p$$

$$\mu_3 = p - 3pp + 2p^3 = p - p^2 - 2p^2 + 2p^3$$

$$= pq - 2p^2q = pq(1-p-p) = pq(q-p)$$

$$\sigma^3 = \sigma.\sigma^2 = pq\sqrt{pq}$$

مثال:

لتكن E تجربة رصد نتيجة ولادة ذكر بوصفها نتيجة لعملية ولادة واحدة.

نعلم أن هذه التجربة هي تجربة برنولية، عندئذٍ بفرض أن A حدث ولادة طفل ذكر فإن:

$$P(A)=\frac{1}{2}\Rightarrow P(\bar{A})=1-p=\frac{1}{2}$$

ليكن X متحول عشوائي يعبر عن حالات ظهور الحدث A (المعبر عن حالة النجاح وهي ولادة طفل ذكر) وبالتالي يكون لـ X جدول التوزيع التالي:

$X=x$	0	1
$f(x)$	$1/2$	$1/2$

إن الدالة $f_X(x)$ تتمتع بالخواص التالية:

1- $f_X(0)=\frac{1}{2}\geq 0\ ,\ f_X(x)=\frac{1}{2}\geq 0$

2- $\sum_{k=1}^{\infty} f_X(x)=\sum_{k=1}^{2} f_X(x)=\frac{1}{2}+\frac{1}{2}=1$

أي أن $f_X(x)$ دالة كثافة احتمالية

الدالة المولدة للعزوم:

$$M_X(t)=q+pe^{t}=\frac{1}{2}(1+e^{t})$$

الدالة المميزة:

$$\varphi_X(t)=q+pe^{it}=\frac{1}{2}(1+e^{it})$$

التوقع الرياضي:

$$E(X)=p=\frac{1}{2}$$

التشتت:

$$\sigma^2(X) = pq = \frac{1}{4}$$

الانحراف المعياري:

$$\sigma(X) = \sqrt{pq} = \frac{1}{2}$$

التناظر:

$$\gamma = \frac{\mu_3}{\sigma^3} = \frac{q-p}{1/2} = 0$$

والتوزيع متناظر.

التوزيع الثنائي

لتكن E تجربة برنولية ما و A حدثاً مرتبطاً بهذه التجربة (يعبر عن حالات النجاح) بحيث أن:

$$P(A) = p \ ,P(\bar{A}) = 1 - p = q$$

$\bar{A}$ الحدث المتمم للحدث A ولنكرر التجربة E وبشكل مستقل تحت نفس الشروط n مرة. بفرض أن X متحول عشوائي يمثل عدد مرات وقوع الحدث A خلال تكرار التجربة n مرة فإن مجموعة قيم X هي $0,1,2,\dots,n$ كما أن دالة احتمال X هي:

$$P_x = P(X = x) = \binom{n}{x} p^x q^{n-x} \quad ,x = 0,1,2,\dots,n$$

عندئذٍ يأخذ جدول توزيع X الشكل التالي:

$X = x$	0	1	2	....	k	....	n
$f(x)$	p_0	p_1	p_2	....	p_k	....	p_n

نقول عن المتحول العشوائي X في هذه الحالة أنه يخضع للتوزيع الثنائي بالوسيطين n و p ونعبر عن ذلك اختصاراً بـ $X \sim b(n,p)$. وسمي كذلك لأن المقدار $\binom{n}{x} p^x q^{n-x}$ هو الحد $x+1$ في منشور ثنائي الحد $(p+p)^n$.

بعض خواص التوزيع الثنائي:

1) $M_X(t) = E(e^{Xt}) = (q + pe^{t})^{n}$

2) $\varphi_X(t) = E(e^{iXt}) = (q + pe^{it})^{n}$

3) $E(X) = \frac{dM_X(0)}{dt} = \left[(npe^{it})(q + pe^{t})^{n-1}\right]_{t=0} = np$

4) $\sigma_X^2 = \alpha_2 - \alpha_1^2 = np + n^2p^2 - np^2 - n^2p^2 = np(1-p) = npq$

5) $\sigma_X = \sqrt{npq}$

6) $\gamma = \frac{\mu_3}{\sigma^3} = \frac{q-p}{\sqrt{npq}}$

وبالتالي إذا كان $p = q$ فإن التوزيع متناظر بالنسبة إلى توقعه الرياضي وإذا كان $p > q$ يكون التوزيع منحرفاً نحو اليسار وإذا كان $p < q$ فإن التوزيع منحرف نحو اليمين.

مثال:

تمتلك إحدى المشافي 8 سيارات إسعاف فإذا كان احتمال انشغال إحدى هذه السيارات في اللحظة t هو 0.2 احسب احتمال:

- انشغال سيارتين في اللحظة t.
- انشغال سيارة واحدة على الأقل في اللحظة t
- ارسم مضلع التوزيع للمتحول العشوائي X الذي يمثل عدد السيارات المشغولة في اللحظة t.

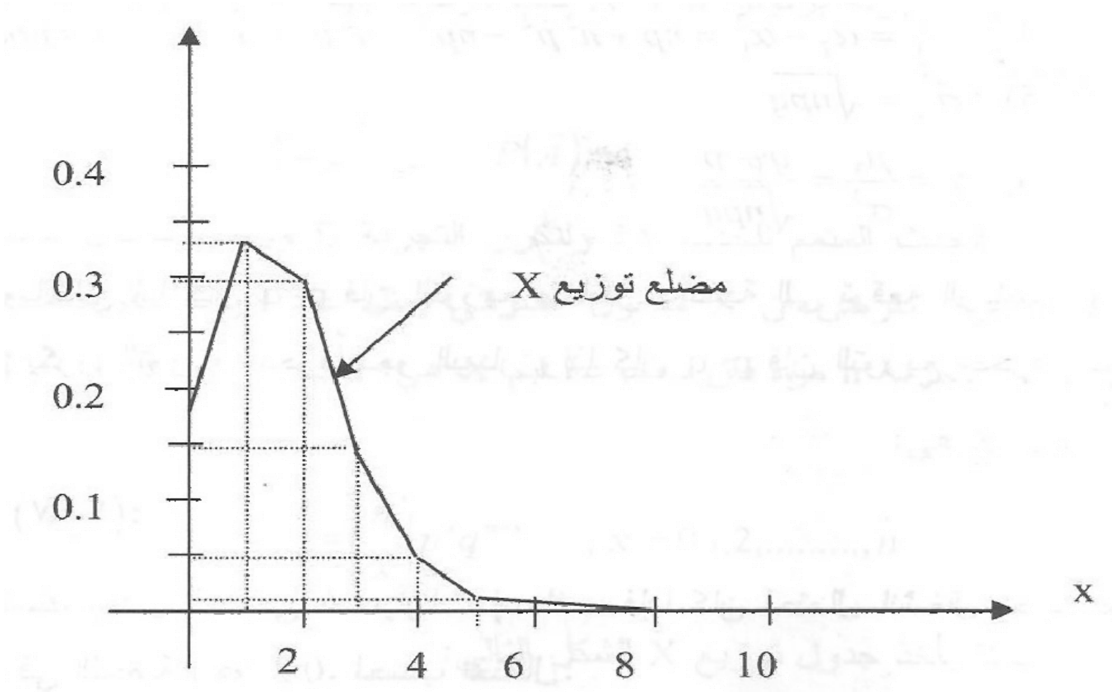
الحل:

إن لـ X التوزيع الثنائي عندئذٍ:

1) $P(X = 2) = \binom{8}{2}(0.2)^2(0.8)^6 \cong 0.294$

2) $P(X \geq 1) = 1 - P(X \leq 1) = 1 - P(X = 0) =$
$1 - \binom{8}{0}(0.2)^0(0.8)^8 \cong 0.83$

$X = x$	0	1	2	3	4	5	6	7	8
$P(X = x)$	0.1678	0.3355	0.2936	0.1468	0.0459	0.0092	0.0011	0.001	0



مثال:

إذا كان احتمال أن يصاب مريض القلب بأزمة قلبية أثناء المعالجة هو 0.05 فإذا كان لدينا 19 مريض بالقلب، والمطلوب:

أ- ما هو احتمال أن تأتي لاثنين منهم على الأقل أزمة قلبية أثناء المعالجة.

ب- ما هو احتمال أن تأتي لاثنين منهم على الأكثر أزمة قلبية أثناء المعالجة.

ج- ما هو العدد المتوقع للمرضى الذين سيصابون بأزمة قلبية أثناء المعالجة.

د- ما هي القيمة الأكثر احتمالاً لعدد المرضى الذين سيصابون بأزمة قلبية أثناء المعالجة.

الحل:

ليكن X متحول عشوائي يرمز لعدد المرضى الذين سيصابون بأزمة قلبية أثناء المعالجة.

أ- إن لـ X التوزيع الثنائي بالوسيطين $n = 19$ و $p = 0.05$ وبالتالي:

$$P(X = x) = C_x^n p^x q^{n-x} \quad ; \quad x = 0,1,2,\dots,19$$

عندئذٍ:

$$P(X \geq 2) = 1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)] \cong 0.2454$$

أو نلاحظ أن $n.p = 0.95 < 5$ وبالتالي يمكننا تقريب التوزيع الثنائي إلى توزيع بواسون بالوسيط $\lambda = n.p = 0.95$ أي:

$$P(X = k) = \frac{(0.95)^k}{k!} e^{-0.95} \quad ; \quad k = 0,1,2,..$$

عندئذٍ:

$$P(X \geq 2) = 1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)] =$$
$$= 1 - 0.3867 - 0.3674 \cong 0.2459$$

بالمقارنة مع الحل السابق نجد أن النتيجة مطابقة حتى ثلاثة أرقام عشرية.

ب-

$$P(X < 2) = P(X = 0) + P(X = 1) + P(X = 2) =$$
$$= C_0^{19}(0.05)^0(0.95)^{19} + C_1^{19}(0.05)^1(0.95)^{18} + C_2^{19}(0.05)^2(0.95)^{17} \cong 0.9334$$

ج-

$$E(X) = n.p = 0.95$$

د- القيمة الأكثر احتمالاً تقع في المجال $[np - q, np + p]$ حيث:

$$np - q = 19(0.05) - 0.95 = 0$$

$$np + p = (19 + 1)(0.05) = 1$$

إذن هناك قيمتان تكون من أجلهما دالة الاحتمال في قيمتها العظمى وهما $X = 0$ و $X = 1$.