# الفصل الرابع

## التوزيع الاحتمالي Probability Distribution

### المتغير العشوائي Random Variable

هو دالة ذات قيمة حقيقية معرفة على مجموعة فضاء العينة حيث S تمثل المنطلق ومستقر الدالة مجموعة جزئية من الاعداد الحقيقية ويرمز لها بإحدى الحروف الكبيرة بحيث ان لكل X عدد حقيقي x قيمة المتغير العشوائي .

**مثال/ رمي قطعة نقود معينة مرتين وكان المتغير X يمثل عدد الصور؟**

الحل/ $S = \{HH, HT, TH, TT\}$

$$X(TT) = 0, X(HT) = X(TH) = 1\ , X(HH) = 2$$

$$x = 0,1,2$$

ويمكن التميز بين نوعين من المتغيرات العشوائية اعتمادا على نوع القيم التي يأخذها المتغير وهي :

١- المتغير العشوائي الذي تختلف قيمة الواحدة عن الاخرى بكميات محددة يسمى المتغير العشوائي المتقطع

٢- المتغير العشوائي الذي يأخذ اي قيمة ضمن حدود معينة يسمى المتغير العشوائي المستمر R.V.C

**التوزيع الاحتمالي المتقطع :-**

هو قانون او جدول يعطي جميع قيم المتغير العشوائي مع جميع الاحتمالات المقترنة مع كل قيمة من قيم المتغير المتقطع ويرمز لـ P بحيث يحقق الشرطين وهما

$$1.P(x) > 0 \quad , 2.P(x_i) = 1$$

وتسمى P(x) دالة احتمالية

**مثال/ بالنسبة للمثال السابق**

احتمال المتغير العشوائي X=0 يعني احتمال عدم الحصول على الصورة

$$P(X=0)=\frac{1}{4}$$

$$P(X=1)=\frac{2}{4}=\frac{1}{2}$$

$$P(X=2)=\frac{1}{4}$$

اذن

$$P(X)=\begin{cases}\frac{1}{4}\ ,x=0,2\\ \frac{1}{2}\ ,x=1\\ 0\ ,0.w\end{cases}$$

او

| X | 0 | 1 | 2 |
|---|---|---|---|
| P(x) | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ |

**مثال/ دالة الكثافة الاحتمالية لرمي الزار للمتغير العشوائي x الذي يمثل عدد النقاط عند رمي الزار ؟**

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

$$=0\qquad ,0.w$$

**التوزيع الاحتمالي المستمر :**

هو التوزيع الذي يأخذ فيه المتغير العشوائي قيما بين حدين ودالته $f(x)$ تكون موجبة لجميع قيم x بين $-\infty < x < \infty$ تسمى $f(x)$ بدالة الكثافة الاحتمالية ومن خواص الدالة

$$1.\ f(x)>0 \qquad 2.\ \int_{-\infty}^{\infty} f(x)=1$$

**التوقع الرياضي :-**

اذا كان x متغير عشوائي بدالة كثافة احتمالية او كتلة احتمالية $p(x), f(x)$ كانت $g(x)$ دالة اخرى بدالة المتغير العشوائي فالتوقع الرياضي لهذه الدالة ويرمز له بـ $E\big(g(x)\big)$ كما في الشكل التالي :

المتغير متقطع $E\big(g(x)\big) = \sum g(x), p(x)$

المتغير المستمر $E\big(g(x)\big) = \int g(x) * f(x)$

**بعض الحالات الخاصة للتوقع الرياضي : -**

أ) **المعدل (Mean) :**

اذا كانت $g(x) = x$ والتوقع الرياضي في هذه الحالة يسمى القيمة المتوقعة للمتغير العشوائي ويرمز له بـ M

ويعني ان القيمة المتوقعة هي في الحقيقة قيمة المتوسط النظري للمجتمع M اي $E(x) = M$

ب) **التباين (Variance)**

اذا كانت الدالة $g(x) = (X - M)^2$ فالتوقع الرياضي في هذه الحالة التباين المتغير العشوائي ويرمز له بـ $S^2 = E(X - M)^2$

**التوزيعات الاحتمالية المتقطعة:**

**توزيع ثنائي الحدين Binomial Distribution**

في حالة تكرار تجربة (n) من المرات و (P) تمثل احتمال النجاح في كل محاولة و q=1−p احتمال الفشل في كل محاولة والمتغير العشوائي (المتقطع), (X) يمثل عدد المرات التي نحصل فيها على النجاح من بين كل المحاولات فان دالة كتلة الاحتمال probability mass function تأخذ الشكل التالي :

$$p(x) = \begin{cases} C_x^n \, P^x \, q^{n-x} & x = 0,1,2, \dots, n \\ 0 & \text{في الاحوال غير المعرفة في اعلاه} \end{cases}$$

الوسط الحسابي لقيم X في هذا التوزيع هو $(\mu_x = np)$

والتباين لقيم X في هذا التوزيع هو $(\sigma_x^2 = npq)$

والانحراف المعياري لقيم X في هذا التوزيع هو $\left(\sigma_x^1 = \sqrt{npq}\right)$

يعتبر توزيع ثنائي الحدين من التوزيعات الاحتمالية المتقطعة عندما يسلك المتغير العشوائي x وفق دالة توزيع ثنائي الحدين نرمز له بـ $x \sim b(n,p)$ حيث كل من $\mathrm{n,p}$ تسمى معلمة parameter التوزيع مثلا $p(x) = C_x^{10}(0.8)^x(0.2)^{10-x}$ تعني ان $p = 0.8\,, n = 10$

**مثال/ نفرض ان $x \sim b(5, 0.9)$ جد الوسط والتباين والانحراف المعياري**

الحل:

$$p(x) = C_x^5 \, (0.9)^x \, (0.1)^{5-x}$$

$$\mu_x = n.p = 5.(0.9) = 4.5$$

$$\sigma_x^2 = n.p.q = 5(0.9)(0.1) = 0.45$$

$$\sigma_x^1 = \sqrt{n.p.q} = \sqrt{0.45}$$

**مثال/ عائلة مكونة من (5) اطفال ما هو احتمال ان يكون فيهم (3) ذكور علما ان نسبة الذكور الى الاناث هي 1:1**

الحل:

$P = \frac{1}{2}\ , q = \frac{1}{2}\ , n = 5\quad , x = 3$

$p(x = 3) = C_3^5 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2$

$= \frac{5!}{3!\ 2!} \left(\frac{1}{8}\right)\left(\frac{1}{4}\right)$

**مثال/ وجد عند مفترق طريقين بأن $\frac{2}{3}$ من السيارات تتجه نحو اليمين و $\frac{1}{3}$ السيارات تتجه نحو اليسار فاذا وصلت (4) سيارات الى المفترق ما هو احتمال اتجاه 3 منها الى اليسار**

الحل:

$p = \frac{1}{3}\quad , q = \frac{2}{3}$

$p(x = 3) = \ C_3^4 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^1$

**مثال/ اذا كان احتمال مريض يشفى من مرض ما $= 0.4$ فاذا دخل المستشفى (5) مرضى مصابين بهذا المرض فما هو احتمال :**

أ) لا يشفى منهم احد

ب) يشفى واحد منهم فقط

ت) يشفى منهم اثنان على الاقل

ث) يشفى منهم اثنان على الاكثر

الحل:

أ) $p = 0.4\quad , q = 0.6\quad , n = 5\quad , x = 0$
$p(x = 0) = C_0^5\ (0.4)^0\ (0.6)^6 = 0.8$

ب) $p = 0.4\ \ , q = 0.6\ \ , n = 5\ \ , x = 1$
$p(x = 1) = \ C_1^5\ (0.4)^1\ (0.6)^4 = 0.26$

ج) $p = 0.4\ , q = 0.6\ \ , n = 5\ \ x \geq 2$

$p(x \geq 2) = p(x = 2) + p(x = 3) + p(x = 4)$

$= C_2^5\ (0.4)^2\ (0.6)^3 + \ C_3^5\ (0.4)^3\ (0.6)^2 + C_4^5\ (0.4)^4\ (0.6)^1$

$+C_5^5\ (0.4)^5\ (0.6)^0$

$= 0.92$

ح) $p(x \leq 2) = p(x = 2) + p(x = 1) + p(x = 0)$

$= C_2^5\ (0.4)^2 (0.6)^3 + C_1^5 (0.4)^1 (0.6)^4 + C_0^5 (0.4)^0 (0.6)^5$

**س/ لوحظ في منطقة معينة ان المطر يهطل خلال عشرة ايام من بين كل ثلاثين يوم . جد احتمال :**

أ) هطول المطر في ثلاثة ايام على الاقل من اسبوع معين

ب) عدم هطول المطر خلال اسبوع معين

**س/ عائلة مؤلفة من (4) اطفال احسب احتمال :**

أ) على الاقل فيها طفل ذكر واحد

ب) على الاكثر (2) ذكر

## التوزيعات المستمرة Continuous Distributions

### التوزيع الطبيعي Normal Distribution

اذا كان المتغير العشوائي X يتوزع توزيعا طبيعيا وله وسط حسابي $\mu$ وتباين $\sigma$ وان معادلة المنحني الطبيعي هي :

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

**خصائص التوزيع الطبيعي :**

- التمثيل البياني له منحني يشبه الجرس ومتماثل بالنسبة للوسط
- يتساوى الوسط والوسيط والمنوال ويقع في المركز
- المنحني متصل
- يقترب المنحني من المحور x ولكنه لا يمسه

- المساحة تحت المنحني تساوي 1 او 100%

يتصف التوزيع الطبيعي الذي وسطه $\mu$ وانحرافه المعياري $\partial$ بالخصائص الاتية :

- يقع 68% تقريبا من البيانات ضمن الفترة $\sigma - \mu \ , \sigma + \mu$
- يقع 95% تقريبا من البيانات ضمن الفترة $2\sigma - \mu \ , 2\sigma + \mu$
- يقع 99% تقريبا من البيانات ضمن الفترة $3\sigma - \mu \ , 3\sigma + \mu$

$\underline{X \approx N} \ (\mu \ , \sigma^2)$

**التوزيع الطبيعي : Normal Distribution**

يقال للمتغير العشوائي X بانه له التوزيع الطبيعي القياسي Standard Normal Distribtuion بوسط حسابي يساوي صفر وتباين يساوي واحد اذا كانت دالة الكثافة الاحتمالية للمتغير العشوائي Z بالصيغة التالية

$F(Z) = \frac{1}{\sqrt{2\pi}} \ e^{-Z^{2t}/2} \ , -\infty < X < \infty$

**مجالات استخدام التوزيع الطبيعي :**

- الاحصاءات البايلوجية كالوفيات والولادات في فترة معينة وفي كل مكان معين
- الظواهر الاجتماعية كالزواج والطلاق تخضع لتوزيع قريب من التوزيع الطبيعي وكذلك نسبة الدخول ومستوى الانتاج
- المقاييس النفسية والعقلية كالذكاء والقابليات العقلية

بما ان $f(x)$ دالة كثافة احتمالية P.D.F لذا فان $\int_{-\infty}^{\infty} f(x)dx = 1$

$X \sim N(0,1)$

والان اصبح من السهل تحويل المتغيرات العشوائية X التي تتوزع توزيعا طبيعيا الى المتغيرات عشوائية Z والتي تتوزع توزيع طبيعي قياسي وذلك بالطريقة التالية :

$Z = \frac{x-\mu}{\sigma}$

**مثلا:** اذا كانت X بين الحدين $x1\ ,\ x2$ فان Z لها ستكون بين $Z1\ ,\ Z2$ حيث ان

$$Z1 = \frac{x1-\mu}{\sigma} \quad , \quad Z2 = \frac{x2-\mu}{\sigma}$$

**ملاحظة :** اذا كان متغير عشوائي له التوزيع الطبيعي فان

$$P(X < x) = N(x)$$

باستخدام جدول التوزيع الطبيعي القياسي

**امثلة :**

- $X \sim N\ (3,4)$ فان

$$f(x) = \frac{1}{\sqrt{2\pi\ 2}}\ e^{-(x-3)^2/2(4)}$$

$$f(x) = \frac{1}{\sqrt{2\pi}}\ e^{-(x-3)^2/8}$$

اذن

- $X \sim N\ (-2,9)$ فان

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-(x+2)^2/18}$$

- $f(x) = \frac{1}{\sqrt{\pi}}\ e^{-(x-3)^2}$ جد $\partial\ ,\mu$

الحل :

$$(x-\mu)^2 = (x-3)\text{^}2 \text{ بما ان}$$

$$\mu = 3 \text{ اذن}$$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} = f(x) = \frac{1}{\sqrt{\pi}} \rightarrow \frac{1}{\sqrt{2\sigma}} = 1$$

$$\rightarrow \sigma = \frac{1}{\sqrt{2}}$$

**مثال: اذا كان الوسط الحسابي لتوزيع طبيعي هو $(\mu = 25)$ والانحراف القياسي (المعياري) له هو $(\sigma = 10)$ جد قيمة $z1 , z2$ بحيث ان :**

$$P(45 < X < 62) = P(z1 < Z < z2)$$

علما بأن : $P(z1 < 0.5) = 0.933 , P(z2 < 12) = 0.8849$

الحل:

$$Z1 = \frac{x1-\mu}{\sigma} = \frac{45-50}{10} = -0.5$$

$$Z2 = \frac{x2-\mu}{\sigma} = \frac{62-50}{10} = 1.2$$

$$P(45 < X < 62) = P(-0.5 < Z < 1.2)$$

$$P(-0.5 < Z < 1.2) = P(Z < 1.2) - P(Z < -0.5)$$

$$= 0.8849 - 0.3085 = 0.5764$$

**مثال:**

**اذا كان X له التوزيع الطبيعي $X \sim N(3,4)$ جد $P(X \leq 2)$ ؟**

الحل:

$$P(X \leq 2) = P\left(\frac{x-\mu}{\sigma} \leq \frac{2-3}{2}\right) = P\left(Z \leq -\frac{1}{2}\right)$$

$$= N\left(-\frac{1}{2}\right) = 1 - (0.5) = 1 - 0.691 = 0.309$$

**مثال:**

**اذا كانت درجات الحرارة خلال شهر نيسان في مدينة ما تتوزع توزيعا طبيعيا بوسط حسابي هو 20C وانحراف معياري 2C جد احتمال ان تكون درجة الحرارة لاحد الايام :**

- اقل من 18C
- اكثر من 22C
- بين 18C و 22C

الحل :

$X - N(20,4)$

- $P(X < 18) = P\left(\frac{x-\mu}{\sigma} < \frac{18-20}{2}\right)$
- $P\left(Z < \frac{-2}{2}\right) = N(-1) = 1 - N(1) = 1 - 0.841 = 0.159$
- $P(X > 22) = 1 - P(X \leq 22) = 1 - P\left(Z \leq \frac{22-20}{2}\right) = 1 - N(1)$
  $= 1 - 0.841 = 0.159$
- $P(18 < X < 22) = P(X < 22) - P(X < 18)$
- $P\left(Z < \frac{22-20}{2}\right) - P\left(Z < \frac{18-20}{2}\right)$
- $= N(1) - N(-1)$
- $N(1) - 1 + N(1) = 2N(1) - 1$
- $2(0.841) - 1 = 0.682$

**ملاحظة :** $F(-X) = 1 - F(X)$

**تمارين**

**١- اذا كان متوسط انتاج الدونم من الذرة الصفراء هو 800 كغم وبانحراف قياسي قدره 40 كغم وعلى فرض ان كمية المحصول يتتبع التوزيع الطبيعي . ما هو احتمال ان نباتا يعطي محصولا بين 834-778 كغم ؟**

**٢- اوجد $X - N(2,25)$**

- $P(x < 0)$
- $P(0 < X < 10)$
- $P(|X| < 10)$

**٣- اذا كان اوزان الف طفل عند الولادة يتوزع طبيعي بوسط حسابي (8.5) وانحراف قياسي (معياري) (0.5) بأون احسب عدد الاطفال بأوزان**

- **اقل من 9.7**
- **بين 7.2 و 8.5**