المحور الثالث: تقارب بعض التوزيعات الاحتمالية

1) التقريب بين قانون ثنائي الحدين وقانون poisson:

بالإضافة إلى استخدامات توزيع بواسون التي سبق ذكرها، فإن لهذا التوزيع استخداما آخر شائعا يتمثل في أنه تقريب جيد لتوزيع ذي الحدين.

$$X \to B(n,p) \longrightarrow X \to P(\lambda)$$

القانون الأصلي القانون المقرب إليه

فلما يكون القانون الأصلي في المسألة هو قانون ثنائي الحدين وتكون الحسابات معقدة نوعا ما، وتسهيلا للحسابات نطبق قانون poisson لأنهما يعطينا تقريبا نفس النتائج.

كما سنلاحظ أن أكبر الاحتمالات المحسوبة في القانون الأصلي رغم صعوبتها مرتبطة بالقيم الصغيرة للمتغير العشوائي X أي أن هذه الظاهرة يمكن اعتبارها ظاهرة نادرة ويستحسن أن نطبق عليها قانون poisson ووجود الجداول الاحتمالية لهذا القانون يسهل علينا الحسابات، لهذا كذلك شروط مضبوطة:

- **<u>الشرط النظري</u>**: أن يكون الاحتمال p في قانون ثنائي الحدين ضعيفا، وبالتالي تعتبر الظاهرة نادرة.
- **<u>الشرط التطبيقي</u>**: ويتمثل في شرطان:
  - حجم العينة كبير $(n \geq 30)$.
  - $n\,p < 5$

<u>مثال</u>:

تظهر الخبرة الماضية لإنتاج مؤسسة معينة أن نسبة 2% من وحدات الإنتاج هي وحدات غير صالحة. للتأكد من ذلك نختبر 30 وحدة منتجة من طرف هذه المؤسسة.

1- ما هو القانون الاحتمالي لهذه المسألة ؟ أحسب $E(X)$، $V(x)$، $\partial(X)$.

2- ما هو احتمال أن يكون عدد الوحدات غير الصالحة ضمن العينة يساوي 1، ما بين(0 و3)، (أكثر من3).

الحل:

1− n=30، نسبة الإنتاج غير الصالح p =0,02، ظاهريا القانون الاحتمالي هو قانون ثنائي الحدين رغم أنه موضوعيا السحب يتم بدون إرجاع.

$$X \rightarrow B(30.0,02)$$

E(X) = n p = 30 . 0,02 = 0,60 مصباح فاسد

V(X) = n p q = 30 0,02 0,98 = 0,588

$\sigma(X) = \sqrt{V(X)} = \sqrt{0{,}588} = 0.77$ مصباح فاسد

نلاحظ أن V(x) = E(X)

2− حساب الاحتمالات:

$$X \in [1,2,3,...........,30]$$

$$P(X=1) = C_{30}^{1}(0{,}02)^{1}(0{,}99)^{30-1}$$

$$P(0 \langle X \langle 3) = P(X+1) + P(X=2)$$
$$= C_{30}^{1}(0{,}02)^{1}(0{,}99)^{30-1} + C_{30}^{2}(0{,}02)^{2}(0{,}99)^{28}$$

P(X> 3) = P(X=4) + P(X=5) +.........+P(X =30)

نلاحظ أن هناك صعوبة كبيرة في إجراء الحسابات وعليه نفضل عملية تقريب إلى قانون poison لأنه يبدو p ضعيف وأن: E(X) =V(X)

التأكد من شروط التقريب:

- $n \geq 30$ محقق.
- $n\,p < 5$، n p = 0,6 وهو محقق

و بالتالي فالشروط محققة ومن ثم يمكن التقريب إلى قانون poisson.

$$X \rightarrow B(30.0,002) \longrightarrow X \rightarrow P(0,6)$$

ونتم حل المسألة باستعمال قانون poisson.

2) التقريب بين قانون فوق الهندسي وقانون ثنائي الحدين:

$$X \to H(N,n,p) \longrightarrow X \to B(n,p)$$

القانون الأصلي القانون المقرب إليه

كما تمت الاشارة إليه سابقا فإنه إذا كانت N كبيرة جدا مقارنة مع n فإنه لا يوجد فرق بين المعاينة بإعادة أو بدون إعادة، وفي الواقع إذا كانت $N \to \infty, N_1 \to \infty$ بحيث يكون $\frac{N_1}{N} \to P$ و يكون $0 \le P \le 1$ فإن:

$$\frac{\binom{x}{N_1}\binom{n-x}{N-N_1}}{\binom{n}{N}} \to \binom{x}{n} P^x (1-P)^{n-x}$$

بشرط أن تكون n صغيرة مقارنة مع N وعليه يصبح:

$$P(X=x) = \binom{x}{n} P^x (1-P)^{n-x}$$

وقد وجد أن هذا التقريب يكون جيدا عندما تكون $N > 50$ و $\frac{N_1}{N} \le 0.1$

فالقانون الأصلي في المسألة هو قانون فوق الهندسي ولكن ولغرض تسهيل الحسابات نستعمل مكانه "قانون ثنائي الحد" خاصة في حساب الاحتمالات. لأنه ثبت تجريبيا أن قانون ثنائي الحدين يعطينا تقريبا نفس النتائج التي يعطيها قانون فوق الهندسي مع ارتياب طفيف يمكن التنازل عنه، لكن تتم عملية التقريب إذا توفرت بشروط مضبوطة.

- ✐ **الشرط النظري العام**: وهو أن يكون حجم العينة "n" صغير جدا مقارنة بحجم المجتمع N.
- ✐ **الشرط التجريبي (التطبيقي)**: هو أن تكون نسبة العينة إلى المجتمع أقل من 5% أي $\frac{n}{N} \le 5\%$.

3) التقريب بين قانون ثنائي الحدين والقانون الطبيعي:

$$X \to B(n,p) \longrightarrow X \to N(\mu,\sigma)$$

أي الانتقال من قانون عشوائي منفصل إلى قانون عشوائي متصل، وعليه كل الاحتمالات تصبح مرتبطة بمجال، واحتمال نقطة يصبح لا يذكر أي $P(X=x) = 0$، ونستعمل جدول القانون الطبيعي المعياري لحساب الاحتمالات.

تتم عملية التقريب بين قانون ثنائي الحدين والقانون الطبيعي لتسهيل عملية الحساب ذلك أنه إذا كانت n حجم العينة كبير تصبح عملية حساب الاحتمالات صعبة، وتتم عملية التقريب وفقا لشروط مضبوطة.

☜ **<u>الشرط النظري:</u>**

- إذا كان n (حجم العينة) كبير جدا
- إذا كان p يقارب 0,5 أي ظاهرة عادية (ليست نادرة).

☜ **<u>الشرط التطبيقي:</u>**

- $n\,p \geq 5$

أما إذا كان P قريبا من 0 الصفر أو 1 الواحد الصحيح، فإن التقريب يكون غير جيد.

نظرية: إذا كانت المتغير العشوائي x يتوزع وفق توزيع ذي الحدين بمعلمتين P و n، وكانت $np \geq 5$ و $n \geq 20$، وكان المتغير العشوائي y يتوزع وفق التوزيع الطبيعي بمتوسط $\mu = n *$ وتباين $\sigma^2 = npq$، فإنه لأي عددين صحيحين a و b يكون.

$$P(a \leq X \leq b) \cong P(a + 0.5 \leq Y \leq b + 0.5)$$

$$= P\left(\frac{(a+0.5)-np}{\sqrt{npq}} \leq \frac{Y-np}{\sqrt{npq}} \leq \frac{(b+0.5)-np}{\sqrt{npq}}\right)$$

$$= P\left(\frac{(a+0.5)-np}{\sqrt{npq}} \leq z \leq \frac{(b+0.5)-np}{\sqrt{npq}}\right)$$

$$= \varphi\left(\frac{(b+0.5)-np}{\sqrt{npq}}\right) - \varphi\left(\frac{(a+0.5)-np}{\sqrt{npq}}\right)$$

حيث أن $\varphi(.)$ تمثل دالة التوزيع التراكمي للتوزيع الطبيعي المعياري.

ونلاحظ ان التوزيع الطبيعي توزيع متصل بينما توزيع ذو الحدين توزيع منفصل، لذلك وجد العالم ياتس Yates أن التقريب يكون أكثر دقة وواقعية، إذا أجرينا تصحيحا على قيمة x وذلك إما بإضافة أو طرح 0.5 منها، ويسمى هذا التصحيح أحيانا بتصحيح ياتس نسبة إلى العالم ياتس، أو معامل التصحيح أو تصحيح الاتصال (Continuity correction)، وعليه فمن المألوف استعمال معامل التصحيح 0.5 الذي يعد ضروريا عند حساب الاحتمال P(X=x).

فمثلا عند حساب احتمال قيمة صحيحة باستعمال التوزيع المتقطع أي ((P(X=x) فهذا ممكن، لكن باستعمال التوزيع الطبيعي المستمر فهذا غير ممكن، لأن الاحتمال هو مساحة ومساحة النقطة تساوي الصفر.

لهذا يعوض احتمال النقطة في التوزيع المستمر بمساحة بين النقطتين (x−0.5) و (x+0.5)، باستعمال معامل تصحيح الاتصال.

ويعوض احتمال $P(X \leq x)$ في التوزيع المتقطع بالاحتمال $P(X \leq x + 0.5)$ في التوزيع الطبيعي المستمر.

ويعوض احتمال $P(X < x)$ في التوزيع المتقطع بالاحتمال $P(X \leq x - 0.5)$ في التوزيع الطبيعي المستمر.

ويعوض احتمال $P(X \geq x)$ في التوزيع المتقطع بالاحتمال $1 - P(X \leq x - 0.5)$ في التوزيع الطبيعي المستمر.

ويعوض احتمال $P(X > x)$ في التوزيع المتقطع بالاحتمال $1 - P(X \leq x + 0.5)$ في التوزيع الطبيعي المستمر.

مثال:

ألقيت قطعة نقود غير متوازنة 15 مرة، ولنفرض أن احتمال ظهور الصورة H هو 0.4، وأن المتغير العشوائي x يمثل عدد مرات ظهور الصورة، احسب الاحتمالين P(X=4)، $P(7 \leq X \leq 9)$، باستخدام دالة الكثافة الاحتمالية لتوزيع ذي الحدين ثم باستعمال دالة الكثافة الاحتمالية للتوزيع الطبيعي.

الحل:

بما ان المتغير العشوائي x يتوزع وفق توزيع ذي الحدين بمعلمتين n=15, p=0.4، فإن دالة الكثافة الاحتمالية تكون:

$$P(X=x)=\begin{cases}\binom{x}{15}0.4^x0.6^{15-x} & x=0,1,\dots.15\\ 0 & O/w\end{cases}$$

$$P(X=4)=0.1268$$

$$P(7\le X\le 9)=P(X\le 9)-P(X\le 6)$$

$$P(7\le X\le 9)=\sum_{x=0}^{9}\binom{x}{15}0.4^x0.6^{15-x}-\sum_{x=0}^{6}\binom{x}{15}0.4^x0.6^{15-x}$$

$$P(7\le X\le 9)=0.9662-0.6098=0.3564$$

باستخدام التقريب للتوزيع الطبيعي نجد أن:

$$\mu=np=15*0.4=6 \qquad \sigma^2=npq=15*0.4*0.6=3.6$$

حساب الاحتمال $P(X=4)$ باستخدام التقريب نجد أن المساحة تحت منحنى التوزيع الطبيعي هي المساحة بين النقطتين $x_1=3.5,\quad x_2=4.5$

$$P(X=4)=P(3.5\le x\le 4.5)=P(-1.316\le z\le -0.789)$$
$$=\varphi(-0.789)-\varphi(-1.316)$$
$$=0.1210$$

وهذه القيمة قريبة جدا من القيمة المحسوبة بواسطة توزيع ذو الحدين.

حساب الاحتمال $P(7\le X\le 9)$ باستخدام التقريب نجد أن المساحة تحت منحنى التوزيع الطبيعي هي المساحة بين النقطتين $x_1=6.5,\quad x_2=9.5$

$$P(7\le X\le 9)=P(6.5\le x\le 9.5)=P(0.263\le z\le 1.482)$$
$$=\varphi(1.482)-\varphi(0.263)$$
$$=0.3636$$

وهذه القيمة قريبة جدا من القيمة المحسوبة بواسطة توزيع ذو الحدين.

4) تقريب توزيع بواسون إلى التوزيع الطبيعي:

يتم تقريب توزيع بواسون إلى التوزيع الطبيعي عندما يتجاوز المقدار $\lambda = np \geq 15$ ويستخدم بعض الاحصائيين التقريب لما يكون $\lambda = np \geq 10$ ، وفي هذه الحالة يتم تقريب توزيع بواسون إلى التوزيع الطبيعي المعياري، بوضع: $\mu = \lambda, \quad \sigma = \sqrt{\lambda}$ وبالتالي يصبح:

$$\lambda > 15 \rightarrow P(X \leq x) = F_{poisson}(x; \lambda) = \varphi(\frac{x + 0.5 - \lambda}{\sqrt{\lambda}})$$

وأيضا في الحالتين التاليتين:

$$P(X \leq x) = F_{poisson}(x; \lambda) \cong \varphi(\frac{x + 0.5 - np}{\sqrt{np}})$$

$$P(X = x) = F_{poisson}(x; \lambda) \cong \varphi\left(\frac{x + 0.5 - np}{\sqrt{np}}\right) - \varphi\left(\frac{x - 0.5 - np}{\sqrt{np}}\right)$$

مثال:

إذا كانت نسبة المعيب من المصابيح المنتجة في مصنع هي 4% ، وسحبنا عينة عشوائية من 400 مصباحا، فالمطلوب هو:

- احسب احتمال أن يكون ربع حجم العينة على الأكثر من الوحدات المعيبة، افتراضا ان المعيب من الانتاج يتبع توزيع بواسون.

الحل:

$$\mu = n * p = 400 * 0.04 = 16 \quad , \qquad P = 0.04 \; , \quad n = 400$$

والمطلوب هو حساب احتمال أن عشر مصابيح على الاكثر تكون معيبة، أي $P(X \leq 10) = ?$

ولو نستخدم توزيع بواسون سيكون علينا حساب الاحتمال بالطريقة التالية:

$$P(X \leq 10) = P(X = 0) + P(X = 1) + P(X = 2) + \cdots \ldots \ldots \ldots . . P(X = 10)$$

أما باستخدام تقريب توزيع بواسون إلى التوزيع الطبيعي فسنجد:

$$P(X \leq 10) = P\left(z \leq \frac{10 + 0.5 - (400 * 0.04)}{\sqrt{400 * 0.04}}\right)$$

$$= P(z \leq -1.37)$$

$$P(X \leq 10) = 0.0853$$

5) تقريب التوزيع فوق الهندسي إلى التوزيع الطبيعي:

يتم تقريب التوزيع فوق الهندسي لمتغير عشوائي متقطع إلى التوزيع الطبيعي لمتغير مستمر، وفق الصيغة التالية:

$$H.G(n; N; N - N_1) \sim N\left(\mu = n\frac{N_1}{N};\ \sigma = \sqrt{n\left(\frac{N_1}{N}\right)\left(1 - \frac{N_1}{N}\right)}\right)$$

ومنه توزيع المتغير العشوائي المعياري يأخذ الشكل:

$$z = \frac{x - n\frac{N_1}{N}}{\sqrt{n\left(\frac{N_1}{N}\right)\left(1 - \frac{N_1}{N}\right)}}$$

ويحسب الاحتمال بالطريقة التالية:

$$P(X \leq x) = P\left(z \leq \frac{x + 0.5 - n\frac{N_1}{N}}{\sqrt{n\left(\frac{N_1}{N}\right)\left(1 - \frac{N_1}{N}\right)}}\right)$$

وبما أن قيم المتغير العشوائي تتبع التوزيع فوق الهندسي وهو متغير متقطع، بينما التوزيع الطبيعي هو توزيع مستمر، فإننا نستعمل معامل تصحيح الاتصال، (continuity correction)، كما رأينا سابقا.

$$P(X \leq x) = \varphi\left(\frac{x + 0,5 - n\frac{N_1}{N}}{\sqrt{n\left(\frac{N_1}{N}\right)\left(1 - \frac{N_1}{N}\right)}}\right)$$

$$P(X < x) = \varphi\left(\frac{x - 0,5 - n\frac{N_1}{N}}{\sqrt{n\left(\frac{N_1}{N}\right)\left(1 - \frac{N_1}{N}\right)}}\right)$$

$$P(X=x)=\varphi\left(\frac{x+0{,}5-n\frac{N_1}{N}}{\sqrt{n\left(\frac{N_1}{N}\right)\left(1-\frac{N_1}{N}\right)}}\right)-\varphi\left(\frac{x-0{,}5-n\frac{N_1}{N}}{\sqrt{n\left(\frac{N_1}{N}\right)\left(1-\frac{N_1}{N}\right)}}\right)$$

$$P(X\geq x)=1-\varphi\left(\frac{x-0{,}5-n\frac{N_1}{N}}{\sqrt{n\left(\frac{N_1}{N}\right)\left(1-\frac{N_1}{N}\right)}}\right)$$

$$P(X>x)=1-\varphi\left(\frac{x+0{,}5-n\frac{N_1}{N}}{\sqrt{n\left(\frac{N_1}{N}\right)\left(1-\frac{N_1}{N}\right)}}\right)$$