جامعة محمد خيضر بسكرة
كلية العلوم الاقتصادية و التجارية
و علوم التسييــر

قسم مجال العلوم الاقتصادية والتسيير
و العلوم التجارية *LMD-SEGC*
السنــة الأولــى

محاضرات في مقياس

الإحصـــاء الرياضي.

المحور الخامس: التوزيعات الاحتمالية الخاصة

الجزء الأول: التوزيعات الاحتمالية الخاصة المتقطعة.

(التوزيع المنتظم، الثنائي، البواسوني، فوق الهندسي).

إعداد الدكتور هاشمي عبابسة.

h.ababsa@univ-biskra.dz

المحور الخامس: التوزيعات الاحتمالية الخاصة

الجزء الأول: التوزيعات الاحتمالية الخاصة المتقطعة.

تمهيد: تهدف نظرية الاحتمالات إلى دراسة التجارب العشوائية من خلال تحديد نموذج احتمالي يمكننا من وصف هذه التجربة وتحليلها بصورة مقبولة.

من الناحية النظرية يمكننا تصورُ عددٍ لانهائي من التجارب العشوائية، وبالتالي تكوينُ عدد لا نهائي أيضا من النماذج الاحتمالية المختلفة، إلا أنه في الواقع هناك عددٌ محدودٌ من التجارب العشوائية التي نصادفها في معظم الأوقات، بحيث تكون لها خصائص مشتركة.

ولهذا سنتطرق في هذه المحاضرة إلى دراسة تلك التجارب العشوائية ونماذجها الاحتمالية المقابلة التي نصادفها بصورة شائعة ومتكررة في التسيير والاقتصاد والتسويق والعلوم الاجتماعية... الخ.

تنقسم قوانين الاحتمالات لهذه التوزيعات الخاصة الى مجموعتين:

- قوانين التوزيعات الاحتمالية المتقطعة.
- قوانين التوزيعات الاحتمالية المستمرة.

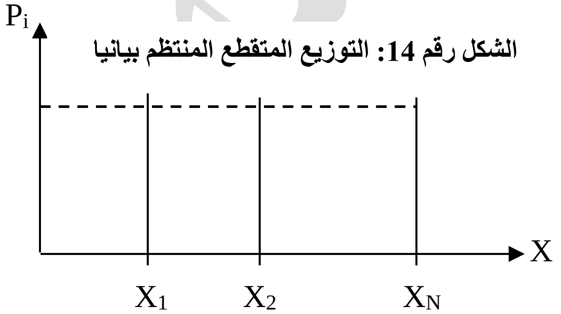
أ- التوزيعات الاحتمالية الخاصة للمتغير المتقطع:

1. التوزيع المنتظم:

- **تعريفه:** نقول إن المتحول العشوائي المتقطع X خاضع للتوزيع المنتظم إذا كان يأخذ القيمَ الممكنة x_i من فضاء الإمكانيات باحتمالات متساوية، أي:

$$P(X = x_1) = P(X = x_2) = \cdots = P(X = x_N)$$

الشكل رقم 14: التوزيع المتقطع المنتظم بيانيا



المصدر: افتراضى

- **تمثيله البياني** : (أنظر الشكل رقم14)
- **خواصه:** ككل توزيع احتمالي:

$* \quad P(x_i) \leq 0 \qquad * \quad \sum P(x_i) = 1$

- **قيمه العددية المميزة:**
  - **التوقع الرياضي:** $E(x) = \sum P.x = \frac{N+1}{2}$
  - **التباين:** $V(x) = E(x^2) - [E(x)]^2 = \frac{N^2-1}{12}$

ملاحظة : العزوم المُمركزة ذات المراتب الفردية لهذا التوزيع كلها معدومة أي :

$\mu_1 = 0 \quad , \quad \mu_3 = 0 \quad , \quad \mu_5 = 0$ ... وهكذا.

مثال01: *رمينا قطعة نرد متوازنة، لنفرض أن X متحول عشوائي يمثل عدد النقاط التي تظهر.*

- *ما هو التوزيع الاحتمالي للمتغير X؟ ولماذا ؟*
- *أحسب توقعه الرياضي وتباينه.*
- *حدد تابع توزيعه.*

الجواب:

- *التوزيع الاحتمالي للمتغير X هو التوزيع هو التوزيع المنتظم، لأن جميع إمكانيات X متساوية الاحتمال.*

- **حساب التوقع الرياضي:**

$$E(x) = \sum_{i=1}^{N} x_i P_i = \frac{21}{6} = \frac{N+1}{2} = \frac{7}{2} = \mathbf{3.5}$$

الجدول رقم 08: حساب التوقع الرياضي.

$P_i . X_i$	P_i	x_i
1/6	1/6	1
2/6	1/6	2
3/6	1/6	3
4/6	1/6	4
5/6	1/6	5
6/6	1/6	6
21/6		

المصدر: معطيات المثال 01.

- **حساب التباين:**

$$V(x) = \frac{N^2 - 1}{12} = \frac{6^2 - 1}{12} = \frac{35}{12} = \mathbf{2.92}$$

- **تحديد تابع التوزيع:**

$$F(x_k) = \begin{cases} 0 \ \left(\text{حدث مستحيل}\right) & x_k < 1 \\ 1/6 & 1 \le x_k < 2 \\ 2/6 & 2 \le x_k < 3 \\ 3/6 & 3 \le x_k < 4 \\ 4/6 & 4 \le x_k < 5 \\ 5/6 & 5 \le x_k < 6 \\ 6/6 = 1 \quad \left(\text{حدث أكيد}\right) & 6 \le x_k < +\infty \end{cases}$$

2. التوزيع الثنائي (الحداني أو ذو الحدَّيْن). ***Distribution Binomiale***

- **تعريفه**: يستند هذا التوزيع الى نموذج " برنولي " (نسبة الى "جميس برنولي " في نهاية القرن 17 م) إذ يستخدم في التجارب ثنائية النتيجة؛ أي التي تؤدي تجربةٌ واحدةٌ منها إلى نتيجتين اثنتين لا غير، مثلا: نجاح ≠ فشل، حياة ≠ موت، إصابة ≠ خطأ، ذكر≠ أنثى ...، ولهذا سُمي هذا التوزيع " توزيع برنولي".
- **قانون احتماله**: نضع " p " يمثل احتمال تحقق إحدى النتيجتين في تجربة واحدة (ونسميه نجاحا) ، ونضع "q" يمثل احتمال عدم تحققها (ونسميه فشلا) ، حيث p + q = 1 .

لنكرر هذه التجربة " N " مرة، إن حصولَنا على " k " نجاحا يعني بالضرورة حصولنا على " N – k " فشلا، وبما أن هذه الـــــN تجربة مستقلةٌ، فإن حادث حصولنا على " k " نجاح يتحقق بعدة طرقٍ عددها C_N^k طريقة.

وبذلك يكون احتمالُ حصولنا على k نجاحا من N تجربة كما يلي:

$$\boxed{P(X = k) = C_N^k p^K q^{N-k}}$$

ويكون المتغير X خاضعا للتوزيع الثنائي ذي المعلمتين N و p، حيث نكتب: $X \backsim B(N,p)$

- **مميزاته العددية** :
  - ✓ **التوقع الرياضي**: $E(x) = N.p$
  - ✓ **التباين**: $V(x) = N.p.q$
- **تابع التوزيع الثنائي**: إذا كان X متحولا عشوائيا خاضعا للتوزيع الثنائي، فإن تابع توزيعه يعطى بالعلاقة الآتية: $F(x_k) = P(x \leq x_k) = \sum_{i=1}^{k} C_N^{x_i}\, P^{x_i} q^{N-x_i}$

كما أن احتمال أي يقع X بين قيمتين M , L مثلا يعطى كما يلي :

$$P(L < X < M) = F(M) - F(L)$$

وبالكيفية نفسها: $P(X > k) = 1 - P(X \leq k) = 1 - F(k)$

__مثال 02__. رمينا قطعة نقود 6 مرات، نفرض أن ظهور الصورة يمثل النجاح.

1. *ما هو احتمال ظهور صورتين؟*
2. *ما هو احتمال ظهور 4 صور على الأقل؟*
3. *ما هو احتمال عدم ظهور الصورة؟*
4. *أحسب μ و $V(x)$ لهذا التوزيع؟*
5. *ما هو احتمال ظهور الصورة أقل من 4 مرات؟*

الجواب: $X \leadsto B(6, 1/2)$

1. $P(X=2) = C_6^2 p^2 q^4 = \frac{6!}{2!(6-2)!}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^4 = \mathbf{\frac{15}{64}}$
2. $P(X \geq 4) = P(X=4) + P(X=5) + P(X=6)$
$$= C_6^4 p^4 q^2 + C_6^5 p^5 q^1 + C_6^6 p^6 q^0$$
$$= \frac{6!}{4!\,(6-4)!}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^2 + \frac{6!}{5!\,(6-5)!}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^1 + \frac{6!}{6!\,(6-6)!}\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^0$$
$$= \frac{15}{64} + \frac{6}{64} + \frac{1}{64} = \mathbf{\frac{22}{64}}$$
3. $P(X=0) = C_6^0 p^0 q^6 = \frac{6!}{0!(6-0)!}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^6 = \mathbf{\frac{1}{64}}$
4. $E(x) = N.p = 6.\left(\frac{1}{2}\right) = \mathbf{3}$
$V(x) = N.p.q = 6\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \left(\frac{6}{4}\right) = \frac{3}{2} = \mathbf{1.5}$
5. $P(X < 4) = 1 - P(X \geq 4) = 1 - \frac{22}{64} = \mathbf{\frac{42}{64}}$

مثال 03: *رمينا زهرة نرد 7 مرات، نفرض أن ظهور 5 أو 6 في أية رمية يمثل النجاح.*

- *ما هو احتمال تحقق النجاح 3 مرات؟*

الجواب: $X \leadsto B(7, 2/6)$

$$P(X=3) = C_7^3 p^3 q^4 = \frac{7!}{3!\,(7-3)!}\left(\frac{2}{6}\right)^3\left(\frac{4}{6}\right)^4 = 35 \times \frac{1}{9} \times \frac{16}{81} = \frac{560}{729} = \mathbf{0.7681}$$

3. توزيع "بواسون": ***Distribution de Poisson***

➢ **تعريفه**: سُمي هذا التوزيع كذلك نسبة إلى عالم الرياضيات والفيزياء الفرنسي "*Siméon Denis Poisson*" الذي اكتشفه عام 1838 م. وعلى غِرار التوزيع الثنائي، فهذا التوزيع يستخدم في التجارب ثنائية النتيجة، إلا أنه يختص بالأحداث " **النادرة الوقوع** " في وحدة قياسية معينة، على سبيل المثال:

- ✓ عدد الأخطاء المطبعية في صفحة من صفحات كتاب.
- ✓ عدد حوادث السيارات خلال أسبوع معين.
- ✓ عدد الحشائش الضارة في المتر المربع من حقل قمح مثلاً.
- ✓ عدد الانتحارات خلال سنة في قُطر من الأقطار.

وفي هذه الحالات جميعا، عادة ما نجد أن N كبيرة جداً، بينما " p " (احتمال النجاح) صغيرا جداً.

➢ **قانون احتماله**: نفرض أن X متحول عشوائي متقطع، نقول أن X خاضع لتوزيع بواسون ذي المعلمة λ إذا كان قانون احتماله كما يلي :

$$\boxed{P(X = x_i) = \frac{\lambda^{x_i}\, e^{-\lambda}}{x_i!}}$$

حيث : λ عدد ثابت يمثل متوسط ظهور النجاحات. (التوقع الرياضي)

ونكتب: $X \leadsto \mathcal{P}(\lambda)$

➢ **مميزاته العددية** :

✓ **التوقع الرياضي**: $E(x) = \lambda = NP$

✓ **التباين** : $V(X) = \lambda$

➢ **العلاقة بين التوزيعين الثنائي و البواسوني**:

إذا كان N كبيرا في التوزيع الثنائي، بينما الاحتمال " p " لظهور حدث ما قريبا من الصفر (أي أن q =1- p قريب من 1) فإننا نعتبر هذا الحدث " نادراً " ، وعمليا سوف نعتبر الحدث نادراً إذا كان عدد الاختبارات على الأقل 50 (N≥50) بينما NP<5 ، في مثل هذه الظروف يكون التوزيع الثنائي قريبا جداً من التوزيع البواسوني الذي يمكن استخدامه كتقريب ممتاز للتوزيع الثنائي.

__مثال 04__: إذا كان 3٪ من مصابيح الكهرباء في أحد المصانع معيبةً، نفرض أننا سحبنا عينة من 100 مصباح، أوجد احتمال أن يكون فيها:

1- مصباح واحد معيب.

2- مصباحان معيبان.

3- أربعة مصابيح معيبة.

4- أعد حساب هذه الاحتمالات بتطبيق التوزيع الثنائي.

5- ماذا تلاحظ.

__الجواب__: *من المعطيات نلاحظ أن:*

- *الظاهرة ثنائية (المصباح إما معيب وإما غير معيب، والنجاح هو الحصول على مصباح معيب).*
- *التوقع الرياضي أقل من 5 ($\lambda = Np = (0.03)100 = 3 < 5$).*
- *حجم العينة أكبر من 50 (N= 100).*

كل ذلك يدعونا لتطبيق التوزيع البواسوني بدلا عن التوزيع الثنائي، حيث $\lambda = 3$، ونكتب $X \leadsto \mathcal{P}(3)$

1- $P(X = 1) = \frac{3^1\, e^{-3}}{1!} = \mathbf{0.1494}$

2- $P(X = 2) = \frac{3^2\, e^{-3}}{2!} = \mathbf{0.2240}$

3- $P(X = 4) = \frac{3^4\, e^{-3}}{4!} = \mathbf{0.1680}$

4- إعادة حساب هذه الاحتمالات بتطبيق التوزيع الثنائي:

- $P(X = 1) = C_{100}^{1} p^1 q^{99} = \frac{100!}{1!(100-1)!}(0.03)^1(0.97)^{99} = \mathbf{0.1471}$
- $P(X = 2) = C_{100}^{2} p^2 q^{98} = \frac{100!}{2!(100-2)!}(0.03)^2(0.97)^{98} = \mathbf{0.2251}$
- $P(X = 4) = C_{100}^{4} p^4 q^{96} = \frac{100!}{4!(100-4)!}(0.03)^4(0.97)^{96} = \mathbf{0.1706}$

5- واضح جليا تقاربُ نتائج التوزيعين، وواضح أيضا أن استخدام التوزيع الثنائي مع الاحداث النادرة أصعب واعقد من استخدام التوزيع البواسوني، ولهذا يُفضل استخدام هذا الأخير في مثل هذه الحالات.

مثال 05: *تتلقى تحويلة الهاتف في الجامعة المكالمات الخارجية بين الساعة 10h والساعة 12h , بمعدل ثلاث مكالمات في الدقيقة.*

المطلوب: *احسب احتمال أن يكون بين الساعة 10.55h و 10.56h :*

1- ولا مكالمة.   2- مكالمة واحدة.   3- على الأقل مكالمتين.   4- على الأكثر مكالمتين

الجواب: *من المعطيات نلاحظ:*

- *أننا أمام ظاهرة ثنائية، النجاحُ فيها هو وصول المكالمة.*
- *التوقع الرياضي أقل من 5 $(\lambda = 3 < 5)$.*
- *N أكبر من 50، فهو غير محدد.*

كل ذلك يدعونا لتطبيق التوزيع البواسوني بدلا عن التوزيع الثنائي، حيث $\lambda = 3$، ونكتب $X \rightsquigarrow \mathcal{P}(3)$

1. $P(X=0) = \frac{3^0\, e^{-3}}{0!} = \mathbf{0.0498}$ *(ولا مكالمة)*
2. $P(X=1) = \frac{3^1\, e^{-3}}{1!} = \mathbf{0.1494}$ *(مكالمة واحدة)*
3. $P(X \geq 2) = 1 - P(X<2) = 1 - [P(X=0) + P(X=1)]$ *(على الأقل مكالمتين)*

$$= 1 - \left(\frac{3^0\, e^{-3}}{0!} + \frac{3^1\, e^{-3}}{1!}\right) = 1 - (0.0498 + 0.1494) = \mathbf{0.8008}$$

4. $P(X \leq 2) = P(X=0) + P(X=1) + P(X=2)$

$$= \frac{3^0\, e^{-3}}{0!} + \frac{3^1\, e^{-3}}{1!} + \frac{3^2\, e^{-3}}{2!} = (0.0498 + 0.1494 + 0.2240 = \mathbf{0.4232}$$

- **تابع توزيع بواسون**: $F(x_k) = P(X \leq x_k) = \sum_{i=0}^{x_k} \frac{\lambda^{x_i}\, e^{-\lambda}}{x_i!}$

ملاحظة:

1- معرفة قيمة λ مسألة ضرورية لحساب الاحتمال وفق قانون بواسون.

2- يمكن استخراج قيمة الاحتمال أو تابع التوزيع البواسوني من خلال جداول إحصائية أُعدت لذلك.

4. التوزيع فوق الهندسي: (التوزيع الهندسي الزائدي) *Distribution hypergéométrique*

- **تعريفه**: عند دراستنا للتوزيع الثنائي يتضح لنا أن عمليات السحب لتكْوين العينة كانت تتم مع إعادة العُنصر المسحوب الى المجتمع، وهذا يترتب عنه أمران، يعتبران بمثابة شرطي تطبيق التوزيع الثنائي، هما:
  - أن جميع السَّحبات (أو التجارب) مستقلة عن بعضها البعض.
  - أن احتمال النجاح في تجربة واحدة p ثبات.

لكن هناك بعض الحالات لا يمكن أن يتم فيها السحب مع الإعادة (كإجراء فحص حول الإصابة بمرض معين) وهنا يختل شرطا تطبيق التوزيع الثنائي (استغلال التجارب، وثبات p) فنلجأ الى تطبيق قانون آخر يسمى قانون فوق الهندسي.

- **قانون احتماله** : إذا كان X متغيرا عشوائيا خاضعا للتوزيع فوق الهندسي، ذي المعالم N، M، n فإننا نكتب $X \backsim H(N, M, n)$، ويعطى قانون احتماله كما يأتي:

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

حيث:

K: عدد النجاحات.

M: عدد العناصر التي يَنْصَبُّ عليها الاهتمام في المجتمع.

N: حجم المجتمع Ω.

n: حجم العينة المسحوبة.

- **مميزاته العددية** :
  - ✓ **التوقع الرياضي**: $E(x) = n.P = n\left(\frac{M}{N}\right)$
  - ✓ **التباين**: $V(x) = \frac{N-n}{N-1}.n.P.q$

وهذا التباين أصغر من تباين التوزيع الثنائي. يسمى المقدار $\frac{\mathrm{N-n}}{\mathrm{N-1}}$ "مؤثرَ عدم الإعادة".

مثال06: *تتكون هيئة التدريس في أحد فروع الجامعة من أربعة كيميائيين وسبعة فيزيائيين، سحبنا -بصورة عشوائية- لجنة مؤلفة من أربعة أساتذة. لنفرض أن X متحول عشوائي يمثل عدد الكيميائيين في هذه اللجنة.*

المطلوب: *حدد التوزيع الاحتمالي للمتغير X.*

الجواب: *نلاحظ أن تجربة سحب هذه اللجنة هي تجربة فوق هندسية، لأن:*

- *التجربة ثنائية النتيجة، فكل أستاذ مسحوب إما أن يكون كيميائيا وإما أن يكون فيزيائيا.*
- *السحب يتم دون إعادة، إذ لا يعقل سحب أستاذ ثم اعادته لأن هذا من شأنه جعل تكرار سحب الأستاذ نفسه حدثا ممكنا، وعلى ذلك فالسحب دون إعادة سيترتب عليه اختلال الشرطين السابقين لتطبيق التوزيع الثنائي.*

وعلى ذلك فإن X (عدد الكيميائيين في اللجنة) خاضع للتوزيع فوق الهندسي $X \backsim H(11,4,4)$، قيمه الممكنة هي:

$\Omega = \{0, 1, 2, 3, 4\}$

$$P(X=0)=\frac{C_4^0 C_{11-4}^{4-0}}{C_{11}^4}=\mathbf{0.106}$$

$$P(X=1)=\frac{C_4^1 C_7^3}{C_{11}^4}=\mathbf{0.424}$$

$$P(X=2)=\frac{C_4^2 C_7^2}{C_{11}^4}=\mathbf{0.380}$$

$$P(X=3)=\frac{C_4^3 C_7^1}{C_{11}^4}=\mathbf{0.085}$$

$$P(X=4)=\frac{C_4^4 C_7^0}{C_{11}^4}=\mathbf{0.003}$$

ومن هنا يمكننا وضع التوزيع الاحتمالي للمتغير X كما يأتي:

الجدول رقم09: التوزيع الاحتمالي للمتغير فوق الهندسي X

x_i	0	1	2	3	4	المجموع[1]
$p(x_i)$	0.106	0.424	0.380	0.085	0.003	**1.000**

المصدر: حلول المثال 06.

[1] قد يلاحظ الطالب أن هذا المجموع لا يساوي واحد بالضبط بل يساوي 0.998 وهذا ببساطة راجع لمشاكل التقريب فقط.