الإحصاء والاحتمالات- المحاضرة الرابعة

Statistics and probabilities- Lecture 4

Dr Fadi KHALIL

Doctor lecturer in statistics and programing

في هذه المحاضرة سيتم التطرق لـ

بعض التوزيعات الإحتمالية المنقطعة

- بيرنوللي Bernoulli،

- الثنائي Binomial،

- بواسون Poisson.

- فوق الهندسي Hypergeometric

يمكن تصنيف التوزيعات الاحتمالية حسب المتغيرات إلى توزيعات احتمالية للمتغيرات المنقطعة توزيعات احتمالية للمتغيرات المستمرة.

9-2-1- التوزيعات الاحتمالية المنقطعة Discrete probability distribution:

9-2-1-1- قانون بيرنولي *Bernoulli*:

بفرض أنّ تجربة ما لها نتيجتين فقط "نجاح: أو "فشل"، ولنفترض أن X متغير عشوائي يأخذ إحدى القيمتين [0،1]، $X=1$ عندما تكون النتيجة "نجاح"، و $X=0$ عندما تكون النتيجة "فشل"، نقول أن المتغير X يتبع توزيع بيرنولي $\mathcal{B}\ (p)$ ذو المعامل $p\ \in]0,1[$ إذا كان :

$$P(X=1)=1-P(X=0)=p$$

على سبيل المثال، يستخدم هذا القانون في الحالات لدراسة احتمال عمل نظام ما. فمثلاً إذا كان احتمال عمل نظام ما هي p فيكون احتمال تعطل النظام هي $1-p$ ويطبق هذا القانون في ألعاب الحظ مثل لعبة الليرات النقدية التي رأيناها سابقاً.

انطلاقاً من مفهوم التوقع (expectation) واعتماده على الاحتمال، يمكن حساب التوقع الرياضي لأي متغير عن طريق مجموع حاصل ضرب كل قيمة للمتغير في الاحتمال المرافق، أي أنّ التوقع الرياضي لمتغير عشوائي يأخذ n قيمة:

$$E(x)=p_1*x_1+\ p_2*x_2+\cdots p_n*x_n$$

$$E(x)=\sum_{i=1}^{n}p_i*x_i$$

وهذا يذكر بقانون الوسط الحسابي لبيانات تكرارية $\bar{x}=\sum_{i=1}^{n}\frac{f_ix_i}{n}$، حيث يتضح أنّ الوسط الحسابي هو شكل من أشكال التوقع، وذلك باعتبار أنّ احتمال المرافق لكل قيمة هو عدد التكرارات مقسوماً على عدد القيم الإجمالي.

وبالعودة لتوزيع بيرنولي ومما سبق يمكن حساب التوقع الرياضي لمتغير يخضع لتوزيع Bernoulli عن طريق:

$$E(x)=1*p(x=1)+0*p(x=0)=p$$

أي أنّ التوقع الرياضي لمتغير عشوائي يخضع لتوزيع بيرنولي يساوي احتمال أن يأخذ المتغير العشوائي القيمة 1.

أمّا تباين (*variance*) متغير يخضع لتوزيع بيرنولي يُستنتج كما يأتي:

رأينا سابقاً أنّ تباين متغير هو الوسط الحسابي لمربعات انحرافات القيم عن وسطها الحسابي أو :

$$V(x)=p(1-p)$$

9-2-1-2- القانون الثنائي *Binomial* :

بفرض أنّه تمّ إجراء n تجربة مستقلة، ونتيجة كل تجربة منها إما "نجاح" مع احتمال p أو "فشل" مع احتمال $1-p$.
وإذا كان X يمثل عدد حالات النجاح التي تحدث خلال n تجربة. بالتالي X هو متغير عشوائي يخضع للتوزيع الثنائي (*Binomial*). حيث يعتبر التوزيع الثنائي امتداداً لتوزيع بيرنولي في الحالات التي يكون فيها إعادة لعدد التجارب.

إذاً نقول أن المتغير X، الذي يأخذ قيمه ضمن المجال المنقطع $\{0,\ 1,\ 2, \dots n\}$، يتبع التوزيع الثنائي $\mathcal{B}\ (n,p)$ ذو المعاملات (n,p) إذا كان:

$$P(X=k)=C_n^k p^k (1-p)^{n-k}, \qquad 0 \le k \le n.$$

حيث، $C_n^k=\frac{n!}{k!(n-k)!}$ يمثل عدد المجموعات المختلفة أو الحالات المختلفة لاختيار k حالة (نجاح مثلاً) من أصل n تجربة. على سبيل المثال بفرض $n=5$ و $k=2$، بالتالي، C_n^k تعبر عن عدد الخيارات التي ينتج فيها "نجاح s" من ضمن 5 تجارب (كل تجربة فيها تحتمل نتيجتين نجاح s أو فشل f):

$$C_5^2=\frac{5!}{2!\,(5-2)!}=\frac{5\times4\times3!}{2!\,(3!)}=\frac{20}{2}=10$$

أي أنّ هناك 10 حالات لظهور مرتين للنجاح s خلال خمسة التجارب:

(s, s, f, f, f) (f, s, s, f, f) (f, f, s, f, s) (s, f, s, f, f) (f, s, f, s, f) (s, f, f, s, f) (f, s, f, f, s) (f, f, f, s, s) (s, f, f, f, s) (f, f, s, s, f)

أي انّه من الممكن أن تظهر حالتي النجاح s في البداية مثل *(s, s, f, f, f)* أو أن تظهر في التجربة الثالثة والخامسة مثل *(f, f, s, f, s)*. وكل حالة من هذه الحالات تملك الاحتمال $p^2(1-p)^{5-2}$ وبالتالي من أجل حساب الاحتمال الدقيق لظهور حالتي نجاح s خلال 5 تجارب يتم ضرب هذا الاحتمال بـ C_5^2.

يستخدم هذا القانون مثلاً لحساب احتمال عدد مرات الحصول على صورة عند رمي قطعة نقدية n مرة حيث ترمز p على احتمال الحصول على صورة أو شعار عند كل رمية.

يُعطى توقع متغير عشوائي يخضع للتوزيع الثنائي بالعلاقة الآتية :

$$E(x)=np$$

أمّا تباين متغير يخضع للتوزيع الثنائي:

$$V(x)=np(1-p)$$

مثال 15:

بفرض قمنا بسحب 3 كرات مع الإعادة من صندوق يحتوي على 7 كرات، 3 صفراء و 4 حمراء، والمطلوب:

1- احسب احتمال أن تكون الثلاث كرات صفراء.

2- احسب احتمال أن تكون هناك كرة صفراء، وكرتين حمراء.

3- احسب احتمال أن تكون هناك كرتين صفراء وكرة حمراء.

4- احسب احتمال أن لا تكون هناك ولا كرة صفراء.

الحل:

من الواضح وجود عدد من التجارب (7)، وبفرض X يمثل عدد الكرات الصفراء المسحوبة يمكن القول أنّ X يخضع للتوزيع الثنائي $\mathcal{B}\ (7,p)$ ، وبفرض أنّ احتمال احتمال أن يتم سحب كرة صفراء في التجربة الواحدة هو $P=\frac{k}{n}=\frac{3}{7}$، بالتالي X يخضع للتوزيع الثنائي $\mathcal{B}\ (7,\frac{3}{7})$

1- احتمال أن تكون الثلاث كرات صفراء:

احتمال أن يتم سحب كرة صفراء هو $P=\frac{k}{n}=\frac{3}{7}$، أمّا احتمال سحب كرة حمراء هو $1-p=1-\frac{3}{7}$ يساوي $\frac{4}{7}$:

$$P(X=3)=C_3^3(\frac{3}{7})^3(1-\frac{3}{7})^{3-3}=0.079$$

C_3^3 تعبر عن صيغة رياضية تسمى توافيق، حيث C_n^k تدل على عدد الحالات لاختيار k عنصر من أصل مجموعة تحتوي على n عنصر، ويعطى بالعلاقة الآتية:

$$C_n^k=\frac{n!}{k!\,(n-k)!}$$

وبالتالي $C_3^3=\frac{3!}{3!(3-3)!}=1$

2- احسب احتمال أن تكون هناك كرة صفراء، وكرتين حمراء.

$$P(X=1)=C_3^1(\frac{3}{7})^1(1-\frac{3}{7})^{3-1}=3*\left(\frac{3}{7}\right)^1\left(1-\frac{3}{7}\right)^{3-1}=0.42$$

$$C_3^1=\frac{3!}{1!\,(3-1)!}=\frac{3!}{1!\,2!}=\frac{3\times 2!}{2!}=3$$

3- احسب احتمال أن تكون هناك كرتين صفراء وكرة حمراء.

$$P(X=2)=C_3^2(\frac{3}{7})^2(1-\frac{3}{7})^{3-2}=3*\left(\frac{3}{7}\right)^2\left(1-\frac{3}{7}\right)^1=0.31$$

$$C_3^2=\frac{3!}{2!\,(3-2)!}=\frac{3!}{2!\,1!}=\frac{3\times 2!}{2!}=3$$

4- احسب احتمال أن لا تكون هناك ولا كرة صفراء

احسب احتمال عدم سحب ولا كرة صفراء، نلاحظ ان هذا الاحتمال هو الاحتمال المكمل لاحتمال أن تكون الكرة هي كرة صفراء:

$$1-[P(X=1)+P(X=2)+P(X=3)]$$

$$1-[0.42+0.31+0.079]=0.191$$

مثال 16:

إذا كان من المعروف أنّ المادة الإسمنتية المصنعة من قبل شركة ما قد تكون تالفة باحتمال قدره 0.1. وإذا كانت الشركة تبيع الإنتاج شكل حزم بحيث تتكون الحزمة الواحدة من 5 أكياس، مع منح كفالة بإرجاع الحزمة واسترداد النقود في حال كان عدد الأكياس التالفة في الحزمة الواحدة أكثر من 1.

المطلوب:

1- ما هو احتمال أن لا يتواجد في الحزمة أي كيس تالف؟
2- ما هواحتمال أت يتواجد في الحزمة كيس تالف واحد؟
3- ما هو احتمال إرجاع الحزمة؟

الحل:

بفرض أنّ X يمثل عدد الأكياس التالفة في الحزمة الواحدة، وباعتبار أنّ عدد الأكياس في الحزمة (5) يمثل عدد التجارب، يمكن القول أنّ X يخضع للتوزيع الثنائي $\mathcal{B}\,(5,0.1)$. بتطبيق قانون التوزيع الثنائي *Binomial* يمكن حساب:

1- احتمال حدث ألا يكون هناك أي كيس تالف:

$$p(X=0)=C_5^0 0.1^0(1-0.1)^{5-0}=1*1*0.9^5\approx 0.59049$$

2- احتمال أن يكون هناك كيس تالف واحد فقط في الحزمة.

$$p(X=1)=C_5^1 0.1^1(1-0.1)^{5-1}=5*0.1^1*0.9^4\approx 0.32805$$

4- احتمال إرجاع الحزمة هو احتمال حدث يكون هناك عدد أكياس تالفة أكثر من 1 $p(X>1)$، وهذا الحدث هو متمم لحدث ألا يكون هناك أي كيس تالف، أو أن يكون هناك كيس تالف واحد فقط في الحزمة. أي $p(X=0)\ or\ p(X=1)$ وبالتالي وفقاً لقاعدة الحدث المتمم:

$$p(X>1)=1-[0.59049+0.32805]\approx 0.08146$$

مثال برؤية أخرى:

إذا كان من المعروف أنّ الأقراص المضغوطة المصنوعة من قبل شركة ما قد تكون تالفة باحتمال قدره 0.01. وإذا كانت الشركة تبيع الأقراص على شكل حزم بحيث تتكون الحزمة الواحدة من 10 أقراص، مع منح كفالة بإرجاع الحزمة واسترداد النقود في حال كان عدد الأقراص التالفة في الحزمة الواحدة أكثر من 1.

المطلوب:

1- ما هو احتمال أن يتم إرجاع الحزمة؟

2- إذا قام شخص بشراء ثلاث حزم، ما هو احتمال أن يتم إرجاع حزمة واحدة فقط.

الحل:

بفرض أنّ X يمثل عدد الأقراص التالفة في الحزمة الواحدة، وباعتبار أنّ عدد الأقراص (10) يمثل عدد التجارب، يمكن القول أنّ X يخضع للتوزيع الثنائي $\mathcal{B}\ (10,0.01)$.

احتمال أن يتم إرجاع الحزمة هو احتمال أن يكون هناك عدد أقراص تالفة أكثر من 1، أي $p(X>1)$:

وهو يشمل احتمال:

$$p(X=2)\ or\ p(X=3)\ or\ p(X=4) or\ p(X=5) or\ p(X=6)$$
$$or\ p(X=7)\ or\ p(X=8)\ or\ p(X=9)\ or\ p(X=10)$$

يمكن حساب كل احتمال من هذه الاحتمالات وبعدها جمعها مع بعضها، ولكن في الواقع هذا الأمر يتطلب بعض الوقت، وبالمقال يوجد حل أسهل وهو البحث عن احتمال الحدث المتمم لحدث أن يكون هناك عدد أقراص تالفة أكثر من 1، وهذا الحدث هو ألا يكون هناك أي قرص تالف، أو أن يكون هنال قرص تالف واحد فقط في الحزمة. أي $p(X=0)\ or\ p(X=1)$ وبالتالي وفقاً لقاعدة الحدث المتمم:

$$p(X>1)=1-[p(X=0)+\ p(X=1)]$$

بتطبيق قانون التوزيع الثنائي *Binomial*:

$$p(X=0)=C_{10}^{0}0.01^{0}(1-0.01)^{10-0}=1*1*0.99^{10}\approx 0.904382$$

$$p(X=1) = C_{10}^{1} 0.01^1 (1-0.01)^{10-1} = 10 * 0.01^1 * 0.99^9 \approx 0.09135$$

$$p(X>1) = 1 - [0.904382 + 0.09135] \approx 0.005$$

2- لنفترض الآن أنّ X هو متغير عشوائي يمثل عدد الحزم التي سيتم إرجاعها من الزبون والذي اشترى 3 حزم، بالتالي X للتوزيع الثنائي $\mathcal{B}\ (3, 0.005)$ حيث 0.005 هو احتمال إرجاع الحزمة والمحسوب من الطلب السابق.

بالتالي احتمال أن يتم ارجاع حزمة واحدة يُحسب كما يأتي:

$$p(X=1) = C_3^1 0.005^1 (1-0.005)^{3-1} \approx 0.015$$

***9-2-1-3- قانون بواسون* Poisson :**

في الحالات التي يكون بها عدد مرات وقوع الحدث كبيرة أو تختلف من تجربة لأخرى أو من فترة لأخرى، يكون من الأنسب استخدام توزيع بواسون. حيث يقوم توزيع بواسون بحساب احتمال وقوع الحدث خلال **فترة زمنية معينة**.

يختلف هذا التوزيع عن التوزيع السابق، بأنّه في حالة التوزيع الثنائي يكون عدد مرات وقع الحدث محدود، مثلا في تجربة إلقاء القطعة النقدية 5 مرات نرى أنّ العدد الإجمالي المتاح للحصول على صورة هو 5 مرات. في حين نلاحظ أنّ عدد المرات لاستقبال رسالة الكترونية خلال ساعة معينة قد يكون كبيراً وغير محدود. هنا توزيع بواسون يمكن استخدمه لحساب احتمال وصول عدد معين من الإيميلات خلال ساعة ما.

ويمكن اعتبار توزيع بواسون هو امتداد للتوزيع الثنائي ولكن عنما يكون عدد التجارب n كبير، بينما يكون الاحتمال p صغير.

نقول أن المتغير X يتبع توزيع بواسون $(\lambda)\ \wp$ ذو المعاملات $(\lambda > 0)$ إذا كان:

$$P(X=k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

يعتبر هذا القانون هو امتداد للقانون الثنائي عندما تكون n تسعى للنهاية $+\infty$ و $n * p \to \lambda$.

بمعنى آخر، عندا إجراء n تجربة مستقلة مع احتمال نجاح p، بالتالي عندما تكون n كبيرة و p يأخذ قيمة صغيرة فإنّ عدد مرات النجاح في التجربة يخضع لتوزيع بواسون بمتوسط λ يساوي $\lambda = \mathrm{np}$

ويستخدم لحساب احتمالات:

- الحوادث النادرة مثل عدد الاتصالات الواردة من إحدى القرى البعيدة على مقسم معين خلال فترة معينة.

- عدد الأخطاء المطبعية في الحروف الموجودة في صفحة أو مجموعة من الصفحات في كتاب ما.

حيث يعتبر احتمال وجود حرف مطبوع بشكل خاطئ هو احتمال صغير في ظل العدد الكبير من الحروف المتضمنة في إحدى الصفحات في الكتاب.

- عدد الأشخاص في المجتمع الذي يعيشن أكثر من 100 سنة.

إن عدد الأشخاص المنتمون لمجتم معين قد يكون عدد كبير واحتمال أن يبغوا من العمر أكثر من 100 سنة هو احتمال صغير.

- عدد أرقام الهواتف الخاطئة التي يتم الاتصال بها في **اليوم**.

- عدد الترانزستورات والتي قد تعطلت في أول يوم من استخدامها.

- عدد العملاء الذين يأتون من منطقة نائية والذين يدخلون مكتب البريد في **يوم محدد**.

يُعطى توقع متغير عشوائي يخضع لتوزيع بواسون بالعلاقة الآتية :

$$E(x) = \lambda$$

أمّا تباين متغير يخضع لتوزيع بواسون:

$$V(x) = \lambda$$

مثال 17:

إذا كان متوسط وصول سيارة إسعاف إلى أحد المشافي هو 3 سيارات، فهما هو احتمال وصول سيارة واحدة خلال ساعة معينة؟

$$P(X = 1) = e^{-3}\frac{3^1}{1!} = 0.15$$

- ما هو احتمال وصول سيارتين خلال ساعة معينة؟

$$P(X = 2) = e^{-3}\frac{3^2}{2!} = 0.22$$

- ما هو احتمال وصول 3 سيارات خلال ساعة معينة؟

$$P(X = 3) = e^{-3}\frac{3^3}{3!} = 0.22$$

- ما هو احتمال وصول 4 سيارات خلال ساعة معينة؟

$$P(X=4)=e^{-3}\frac{3^4}{4!}=0.17$$

- ما هو احتمال عدم وصول ولا سيارة خلال ساعة معينة؟

$$P(X=0)=e^{-3}\frac{3^0}{0!}=0.05$$

- ما هو احتمال وصول 4 سيارات على الأكثر، وهو احتمال عد وصول ولا سيارة إسعاف، أو سيارة إسعاف واحدة، أو اثنتين، او ثلاثة، أو أربعة.

$$P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)$$

$$0.05+0.15+0.22+0.22+0.17=0.81$$

ملاحظة:

يعتبر توزيع بواسون *Poisson* هو تقريب للتوزيع الثنائي عندما يكون عدد التجارب n كبير (أكبر أو يساوي 20) واحتمال وقوع الحدث عند إجراء التجربة لمرة واحدة $p<0.05$، ودقة التقريب تزداد طبعا كلمّا ازداد عدد التجارب n . أي عند النظر لتابع الكتلة الاحتمالي في التوزيع الثنائي:

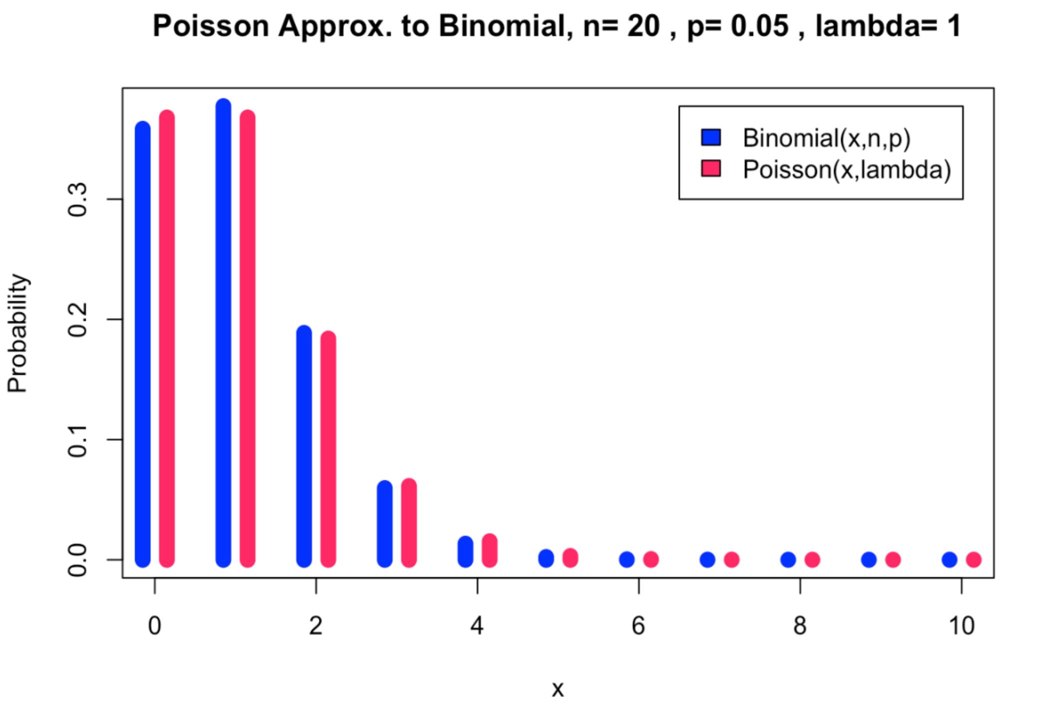
$$P(X=k)=C_n^k p^k(1-p)^{n-k}, \qquad 0\le k\le n$$

يمكن القول أنّه بتطبيق قواعد النهايات أنّ هذه العلاقة تقترب باتجاه صيغة توزيع بواسون

$$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$$

عندما n تقترب باتجاه اللانهاية.

فمثلاً يلاحظ في الشكل الآتي أنّ التوزيع الاحتمالي لعدد مرات النجاح في 20 تجربة وباحتمال قدره 0.05 هو توزيع متشابه (تقريباً) سواء تمّ استخدام قانون الاحتمال لتوزيع بواسون أو الثنائي. وبالطبع دقة التشابه تزداد كلّما ازدادت قيمة n أو انخفضت قيمة p.



مثال *18*:

بفرض أن احتمال أن يوجد عيب في إنتاج إحدى المواد المصنوعة بآلة معينة هو 0.001. فما هو احتمال ان تحتوي عينة مؤلفة من 1000 على عيب واحد على الأكثر.

الحل :

إنّ متوسط عدد العيوب في 1000 قطعة هو $\lambda = \mathrm{np} = 1000 * 0.001 = 1$

واحتمال وجود عيب واحد على الأكثر في الإنتاج $\mathrm{p}(\mathrm{X} \leq 1)$ هو :

$$\mathrm{p}(\mathrm{X} \leq 1) = p(X = 0) + p(X = 1)$$

وبما إنّ عدد التجارب يعتبر كبير (1000) واحتمال وجود عيب في الإنتاج هو صغير 0.001 بالتالي إن المتغير X يخضع لتوزيع بواسون *poisson*:

$$\mathrm{p}(\mathrm{X} \leq 1) = e^{-1}\frac{1^0}{0!} + e^{-1}\frac{1^1}{1!} \approx 0.7358$$

مثال 19:

بفرض أنّ متوسط عدد الحوادث المسجلة أسبوعياً في إحدى الطرق السريعة يساوي 3، فما هو احتمال وقوع حادث واحد على الأقل خلال هذا الأسبوع؟

ليكن X متغير عشوائي تعبر عن عدد الحوادث في إحدى الطرق السريعة هذا **الأسبوع**. ونظراً لأنّ هناك عدد كبير من السيارات التي تعبر الطريق السريع، وأن هناك احتمال صغير لأن تقع إحداها في حادث، فإنّ عدد الحوادث X يخضع لتوزيع بواسون بمتوسط $\lambda = 3$.

احتمال وقوع حادث واحد على الأقل $\mathrm{p}(\mathrm{X} \geq 1)$ هو حدث متمم لحدث عدم وقع أي حادث $\mathrm{p}(\mathrm{X} = 0)$، بالتالي:

$$\mathrm{p}(\mathrm{X} \geq 1) = 1 - \mathrm{p}(\mathrm{X} = 0)$$

$$= 1 - e^{-3}\frac{3^0}{0!} \approx 0.9502$$

4-1-2-9- التوزيع فوق الهندسي The hypergeometric distribution:

في التوزيع الثنائي binomial distribution أنّ للتجربة نتيجتين فقط، وأنّ حساب احتمال عدد مرات النجاح في حال كانت التجربة تتم (مع الإعادة، تذكر مثال الكرات في الصندوق)، وفي حال كانت التجربة تتم من دون إعادة (without replacement)، فإنّ التجارب لم تعد مستقلة عن بعضها، وبالتالي عدد مرات النجاح لم يعد يخضع للتوزيع الثنائي، بل للتوزيع فوق الهندسي (hyper geometric distribution).

للتوضيح أكثر لنورد **المثال الآتي**:

ليكن صندوق يحتوي $N + M$ بطارية، منها N بطارية عاطلة، و M سليمة. إذا تمّ سحب عينة عشوائية (من دون إعادة) بحجم n بطارية، فإنّ عدد العينات المختلفة (بحجم n) التي يتم سحبها يعادل C_{N+M}^{n}.

وبفرض مثلاً أنّ عدد البطاريات السليمة في الصندوق هو $M = 800$ والبطاريات العاطلة هي $N = 50$، فإذا تمّ سحب عينة من بطاريتين $n = 2$ من دون إعادة فما هو احتمال أن لا يوجد أي بطارية عاطلة؟

ليُعرّف X بأنّه عدد البطاريات غير السليمة في العينة العشوائية المسحوبة، بالتالي:

$$p(X = 0) = \frac{800}{850} * \frac{799}{849} = 0.886$$

ما هو احتمال أن يوجد بطارية واحدة عاطلة؟

هنا يوجد حالتين الأولى أن تكون البطارية العاطلة هي الأولى والبطارية السليمة هي الثانية، اما الحالة الثانية فهي العكس، لذلك الاحتمال يُعطى كالآتي:

$$p(X=1)=\frac{50}{850}*\frac{800}{849}+\frac{800}{850}*\frac{50}{849}=0.111$$

ما هو احتمال أن تكون البطاريتين عاطلتين؟

$$p(X=2)=\frac{50}{850}*\frac{49}{849}=0.003$$

يلاحظ في حساب الاحتمالات السابقة أن التجارب غير مستقلة لأنّ السحب من دون إعادة، وبالتالي يتم تخفيض عدد الحالات في كل مرة يتم فيها السحب. ولتبسيط الحسابات يمكن إعطاء الصيغة الآتية:

$$p(X=i)=\frac{C_N^i C_M^{n-i}}{C_{N+M}^n}, \qquad i=0,1,\ldots..\min(N,n)$$

بمعنى آنّ أي متغير عشوائي توزيعه الاحتمالي يعطى بالعلاقة السابقة يطلق عليه المتغير فوق الهندسي *hypergeometric* مع المعاملات N,M,n.

وللتحقق من صحّة هذه الصيغة سيتم استخدمها لحساب الإحتمالات السابقة:

$$p(X=0)=\frac{C_{50}^0 C_{800}^{2-0}}{C_{50+800}^2}=\frac{319600}{360825}=0.886$$

$$p(X=1)=\frac{C_{50}^1 C_{800}^{2-1}}{C_{50+800}^2}=\frac{40000}{360825}=0.111$$

$$p(X=2)=\frac{C_{50}^2 C_{800}^{2-2}}{C_{50+800}^2}=\frac{1225}{360825}=0.003$$

مثال 20:

من المعروف في الهندسة أنّ أي نظام (system) يتكون من مجموعة من المكونات (components) والتي قد تكون زر في واجهة ما، أو نقطة توصيل في إحدى الشبكات .... وبغرض تشكيل أحد الأنظمة المؤلف من 6 مكونات تمّ سحب عينة عشوائية (بحجم 6) من صندوق يحتوي على 20 مكّون (components)، فإذا كان من المعروف أنّ النظام يعمل إذا كان على الأقل 4 من أصل 6 من المكوّنات المسحوبة تتمتع بحالة جيدّة. وكان 15 من أصل 20 مكوّن في الصندوق تتمتع بحالة جيدّة، فما هو احتمال أن يعمل النظام بصورة جيدّة؟

من المعطيات يمكن الاستنتاج أنّ النظام يعمل بصورة جيدّة إذا كان 4 أو أكثر من أصل مكوناته الستة تتمتع بحالة جيدّة، ويلاحظ أيضاً أن السحب بدون إعادة (أي يتم سحب مكوّن ومن ثم يتم وضعه في النظام) بالتالي احتمال أن يعمل النظام بصورة جيدّة هو:

على افتراض X متغير عشوائي يمثل عدد المكوّنات التي تتمتع بحالة جيّدة، فإنّ

$$p(X \geq 4) = p(X = 4) + p(X = 5) + p(X = 6)$$

بتطبيق قانون التوزيع فوق الهندسي السابق:

$$p(X \geq 4) = \frac{C_{15}^{4} C_{5}^{2}}{C_{20}^{6}} + \frac{C_{15}^{5} C_{5}^{1}}{C_{20}^{6}} + \frac{C_{15}^{6} C_{5}^{0}}{C_{20}^{6}} \approx 0.8687$$

يُعطى توقع متغير عشوائي X_i يخضع للتوزيع فوق الهندسي (*geometric*) بالعلاقة الآتية:

$$E(X) = np$$

بحيث، $p = \frac{N}{N+M}$

ويُعطى تباين (*variance*) متغير عشوائي X_i يخضع للتوزيع فوق الهندسي (*geometric*) بالعلاقة الآتية:

$$Var(X) = np(1-p)\left[\frac{N+M-n}{N+M-1}\right]$$

مثال 21:

مجموعة من المكونات الإلكترونية تحتوي على 100 من رقائق الدارات (*circuit board*) المصنّعة محلياً، و200 من رقائق الدارات المستوردة. فإذا تمّ اختيار 4 رقائق بشكل عشوائي ومن دون إعادة، ما هو احتمال أن تكون الرقائق الأربعة مصنّعة محلياً؟

بفرض X_i متغير عشوائي يعبر عن رقائق الدارات المصنعة محلياً

وأنّ N هو عدد الدارات المصنعة محلياً، M عدد الدارات المستوردة.

يلاحظ من معطيات المثال أنّ المتغير يخضع للتوزيع فوق الهندسي لأنّ نتيجة التجربة هي حالتين فقط، والتجارب غير مستقلة. بالتالي لحساب الإحتمال:

$$p(X = 4) = \frac{C_{100}^{4} C_{200}^{4-4}}{C_{100+200}^{4}} = 0.0119$$

ما هو احتمال أن يوجد على الأقل دارتين مصنعتين محليّاً من ضمن الدارات الأربعة:

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

$$p(X \geq 2) = \frac{C_{100}^{2} C_{200}^{2}}{C_{100+200}^{4}} + \frac{C_{100}^{3} C_{200}^{4-3}}{C_{100+200}^{4}} + \frac{C_{100}^{4} C_{200}^{4-4}}{C_{100+200}^{4}}$$

$$p(X \geq 2) = 0.298 + 0.098 + 0.0119 = 0.407$$

ما هو احتمال أن يكون على الأقل دارة واحدة على الأقل مصنعة محلياً،

يمكن حساب هذا الاحتمال على طريقة المثال السابق أو عن طريق الحدث المتمم، فالحدث المتمم لوجود مكوّن واحد على الأقل مصنع محلياً هو حدث عدم وجود أي مكّون مصنع محلياً $(X = 0)$ بالتالي:

$$p(X \geq 1) = 1 - p(X = 0)$$

$$p(X \geq 1) = 1 - \frac{C_{100}^{0} C_{200}^{400-0}}{C_{100+200}^{4}} = 0.804$$

ما هو التوقع الرياضي $E(X)$ و التباين $Var(X)$ لعدد الدارات المصنعة محلياً:

$$E(X) = np$$

بحيث،:

$$p = \frac{N}{N+M} = \frac{100}{100+200} = 0.333$$

$$E(X) = 4 * 0.33 = 1.33$$

التباين $Var(X)$:

$$Var(X) = np(1-p)\left[\frac{N+M-n}{N+M-1}\right]$$

$$Var(X) = 4 * 0.33(1 - 0.33)\left[\frac{100+200-4}{100+200-1}\right] = 0.88$$