Republic of Iraq

Ministry of Higher Education & Scientific Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

محاضرات الاحصاء ١

مدرس المادة : الاستاذ المساعد الدكتور
فراس شاكر محمود

References:

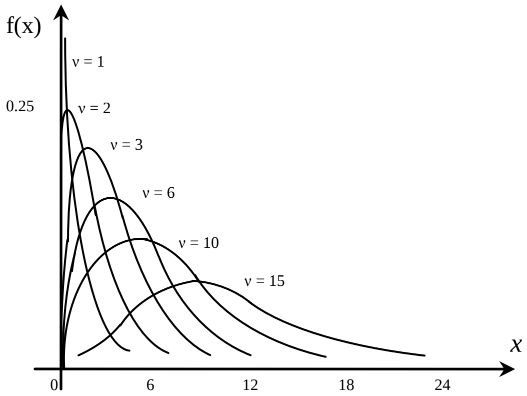
1) **Introduction to Mathematical Statistics, R. V. Hogg and A.T. Craig, (4, 5, 6) edition.**
2) **Mathematical Statistics with Applications, K. M. Ramachandran and C. P. Tsokos, 2009.**
3) **Probability and Statistical Inference, Robert V. Hogg ,Elliot A. Tanis and ale L. Zimmerman, Ninth Edition, Pearson Education, USA,2015.**
4) **Probability and mathematical Statistics, Prasanna Sahoo, University of Louisville,, USA, 2008.**
5) **امير حنا هرمز، الاحصاء الرياضي، مطبوعات جامعة الموصل، ١٩٩٠**
6) **صفاء يونس الصفاوي، الاحصاء، مطبوعات جامعة الموصل، ٢٠٠٨**

الفصل الاول

بعض المواضيع في الاحتمالات :

الدوال غير الخطية: مربع كاي ، فيشر وستيودنت

توزيع مربع كاي Chi – Square Distribution



رسم 1 تدرج منحنى مربع كاي حسب درجة الحرية

توزيع مربع كاي هو من أكثر التوزيعات استخداما في مجال اختبار الفروض بأنواعها، ويمكن تعريفه كما يلي:

لتكن $X_1, X_2, \dots X_\nu$ ، متغيرات عشوائية مستقلة كل منها تتبع التوزيع الطبيعي المعياري $(\mu = 0, \sigma = 1)$. المتغير

$$X = X_1^2 + X_2^2 + \dots + X_\nu^2$$

لها دالة الكثافة التالية:

$$f(x) = \begin{cases} \dfrac{x^{(\nu/2)-1} e^{-x/2}}{2^{\nu/2}\Gamma(\nu/2)} & x > 0 \\ 0 & x \le 0 \end{cases}$$

حيث $\Gamma(\alpha)$ هي الدالة قاما:

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt \qquad \alpha > 0$$

و نقول أن X تتبع التوزيع مربع كاي ب ν درجة حرية ونكتب $X \sim \chi^2_\nu$.

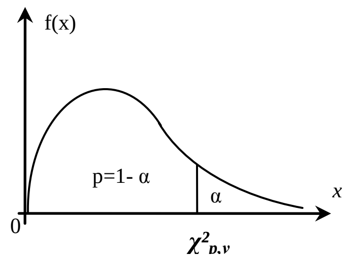
الدالة التجميعية $F(\chi^2)$ تكتب كما يلي:

$$P(X^2 \le x) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^x u^{(\nu/2)-1} e^{-u/2}\, du & x \ge 0 \\ 0 & x < 0 \end{cases}$$

خصائص توزيع مربع كاي $\quad$ **$E(X) = \nu, \quad V(X) = 2\nu, \quad M(t) = (1-2t)^{-\nu/2}$**

دالة التوزيع مربع كاي هي حالة خاصة من توزيع قاما بوضع $\alpha = \nu/2, \beta = 2$. ويأخذ منحنى f(x) شكله حسب قيمة الثابت ν ونلاحظ من الرسم أن المنحنى يبتعد شيئا فشيئا عن المحور العمودي ويأخذ شكلا جرسيا كلما زادت قيمة ν . ونبرهن أنه عند ν كبير $(\nu \ge 30)$ فإن $\sqrt{2\chi^2} - \sqrt{2\nu - 1}$ تتبع التوزيع الطبيعي المعياري.

- في الجداول الاحصائية، تعين نقطة (قيمة المتغير) مربع كاي على المحور الأفقي (أنظر الرسم المقابل) من خلال ν بالإضافة إلى المساحة p على يسار مربع كاي تحت المنحنى $(p = P(X \le \chi^2_{\nu;p})$. وأحيانا تحدد النقطة مربع كاي بدلالة المساحة على يمينها $(\alpha = 1 - p)$ لذلك نجد في كتب الاحصاء كل من الكتابتين: $\chi^2_{p,\nu}$ و $\chi^2_{\alpha,\nu}$



رسم 2 تعيين نقطة مربع كاي على المحور من خلال قيمة p

نظرية: لتكن م ع مستقلة عددها n حيث $X_1 \sim \chi^2_{v_1}, \ldots, X_n \sim \chi^2_{v_n}$ مجموع هذه المتغيرات $X_T = \sum_{i=1}^{n} X_i \sim \chi^2_{v_T}$، $v_T = \sum v_i$.

توزيع ستيودنت Distribution of Student

لتكن المتغيران العشوائيتان المستقلان Y وZ حيث $Y \sim N(0, 1)$ و $Z \sim \chi^2_v$ ؛ المتغير $T = \frac{Y}{\sqrt{Z/v}}$ تتبع توزيع لها دالة الكثافة التالية

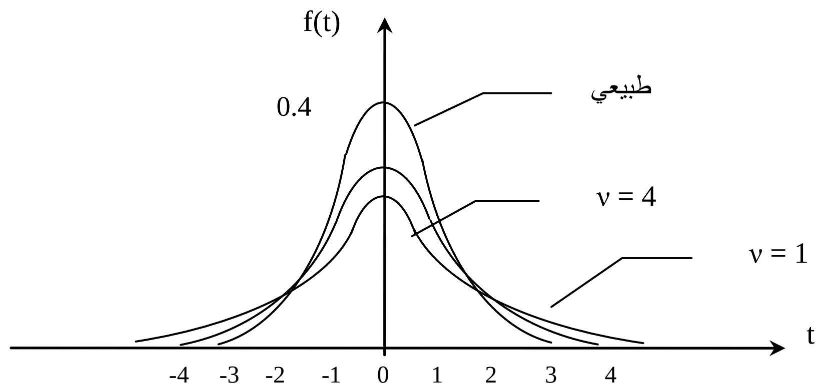
$$\boxed{f(t) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\;\Gamma\left(\frac{v}{2}\right)}\left(1+\frac{t^2}{v}\right)^{-\left(\frac{v+1}{2}\right)} \qquad -\infty < t < \infty}$$

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt \qquad \alpha > 0$$

و نقول أن المتغيرX تتبع توزيع ستيودنت ب v درجة حرية ونكتب: $T \sim t_v$

خصائص توزيع ستيودنت

$$E(T) = 0, \qquad V(T) = v/(v-2) \quad \text{for} \quad (v > 2)$$

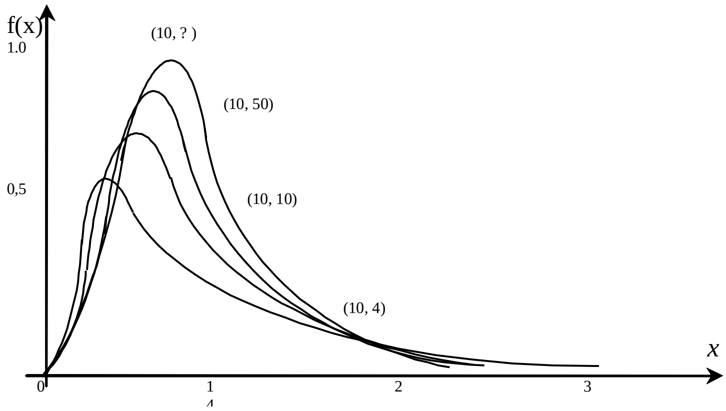


رسم3 تدرج منحنى ستيودنت حسب درجة الحرية

- نلاحظ أن منحنى t متماثل حول المتوسط ، مما يعني أن لكل نقطة موجبة t نقطة مناظرة لها سالبة حيث المساحة تحت المنحنى على يمين t تساوي المساحة تحت المنحنى علي يسار (–t)، ونكتب $t_{1-p} = -t_p$.

- بالإضافة إلى ذلك فإن منحنى f(t) يقترب من المنحنى الطبيعي المعياري كلما زادت قيمة v . وعموما، يعتبر الإحصائيون أن المنحنيان يتطابقان تقريبا عند $v \geq 30$.
- في الجداول الاحصائية، تعين نقطة (قيمة المتغير) t من خلال v والمساحة p على يسار t تحت المنحنى ($p = P(T \leq t_{v;p})$). وأحيانا تحدد النقطة t بدلالة المساحة على يمينها $(\alpha = 1 - p)$ ونكتب : $t_{p,v}$ أو $t_{\alpha,v}$.

توزيع فيشر (F) Distribution of Fisher



رسم 4 تدرج منحنى فيشر حسب درجة الحرية

ليكن لدينا المتغيران العشوائيتان المستقلان $X_1 \sim \chi^2_{v_1}$ و $X_2 \sim \chi^2_{v_2}$. المتغير: $X = \dfrac{X_1/v_1}{X_2/v_2}$ لها دالة الكثافة:

$$f(x) = \begin{cases} \dfrac{\Gamma\left(\dfrac{v_1+v_2}{2}\right)}{\Gamma\left(\dfrac{v_1}{2}\right)\Gamma\left(\dfrac{v_2}{2}\right)} \; v_1^{\frac{v_1}{2}} \; v_2^{\frac{v_2}{2}} \; x^{(\frac{v_1}{2})-1} \; (v_2 + v_1 x)^{-\frac{v_1+v_2}{2}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

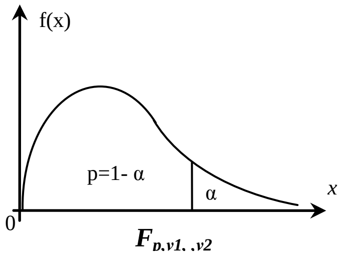
و نقول أن المتغيرX تتبع توزيع فيشر ب v_1 وv_2 درجة حرية ونكتب: $X \sim F_{v1,\, v2}$

خصائص توزيع فيشر:

$$\mu = \frac{v_2}{v_2 - 2} \quad (v_2 > 2) \quad , \quad \sigma^2 = \frac{2v_2^2(v_1 + v_2 - 2)}{v_1(v_2-4)(v_2-2)^2} \quad (v_2 > 4)$$

ويظهر من المعادلة تبعية منحنى f(x) بالإضافة ل x إلى كل من v_1 وv_2 ولذلك تحدد أي نقطة F من خلال ثلاثة معالم: v_1 و v_2 و p, (المساحة تحت المنحنى على يسار النقطة F) ، ونكتب $F_{p,\, v_1,\, v_2}$

وفي الغالب تعطي الجداول الإحصائية قيم F عند $p = 0.95$ و$p = 0.99$.



رسم5 تعيين قيمة F يتم في الجدول يتم من خلال v_1, ,v_2 و P

نظرية ١.

$$F_{1-p,\nu_1,\nu_2} = 1/F_{p,\nu_2,\nu_1}$$

نظرية ٢.

$$F_{1-p,\,1,\,\nu} = t^2_{1-(p/2),\nu}$$

نظرية ٣.

$$F_{p,\nu,\,\infty} = \frac{\chi^2_{p,\nu}}{\nu}$$

خلاصة

يمكن تلخيص أهم ما تضمنه هذا المبحث في الجدول التالي:

أهم ما يجب معرفته عن دالة الكثافة	المتغير العشوائي	التوزيع
$f(x) = 0 \quad \text{for} \quad x \le 0$ $E(X) = \nu, \quad V(X) = 2\nu$	إذا كانت Xi متغيرات عشوائية مستقلة كل منها تتبع التوزيع الطبيعي المعياري، و $X = X_1^2 + X_2^2 + \dots + X_\nu^2$ إذن: $X \sim \chi^2_\nu$	**توزيع مربع كاي** $X \sim \chi^2_\nu$
$E(T) = 0,$ $V(T) = \nu/(\nu-2) \quad \text{for} \quad (\nu > 2)$	لتكن المتغيران العشوائيتان المستقلان Y وZ حيث $Y \sim N(0, 1)$ و $Z \sim \chi_\nu^2$ ؛ إذن: $T = \frac{Y}{\sqrt{Z/\nu}} \sim t_\nu$	**توزيع ستيودنت** $T \sim t_\nu$
$f(x) = 0 \quad \text{for} \quad x \le 0$ $F_{p,\nu,\,\infty} = \frac{\chi^2_{p,\nu}}{\nu}$, $F_{1-p,\,1,\,\nu} = t^2_{1-(p/2),\nu}$; $F_{1-p,\nu_1,\nu_2} = 1/F_{p,\nu_2,\nu_1}$	إذا كانت لدينا متغيرتان عشوائيتان مستقلان حيث: $X_1 \sim \chi_{\nu_1}{}^2$ و $X_2 \sim \chi_{\nu_2}{}^2$ ، فإن $X = \frac{X_1/\nu_1}{X_2/\nu_2} \sim F_{\nu1,\,\nu2}$	**توزيع فيشر** $X \sim F_{\nu1,\,\nu2}$

السلوك التقاربي لبعض التوزيعات الاحتمالية

نتناول في هذا المبحث بعض حالات التقارب الذي يحصل بين عدد من التوزيعات الاحتمالية الشهيرة. ونقصد بالتقارب بين توزيعين (الثنائي و بواسون مثلا) أن يعطي التوزيعان نتائج متقاربة بخصوص احتمال معين، مما يعني إمكانية استخدام توزيعين احتماليين (وأحيانا أكثر) لحساب احتمال معين. علما أننا قد تطرقنا من قبل بإيجاز إلى هذا المفهوم عند دراستنا لهذه التوزيعات.

التقارب بين التوزيع الثنائي والتوزيع الطبيعي

لندرس السلوك التقاربي للمتغير التوزيع الثنائي $X \sim B(n,p)$ عندما تؤول n الأعداد كبيرة جدا.

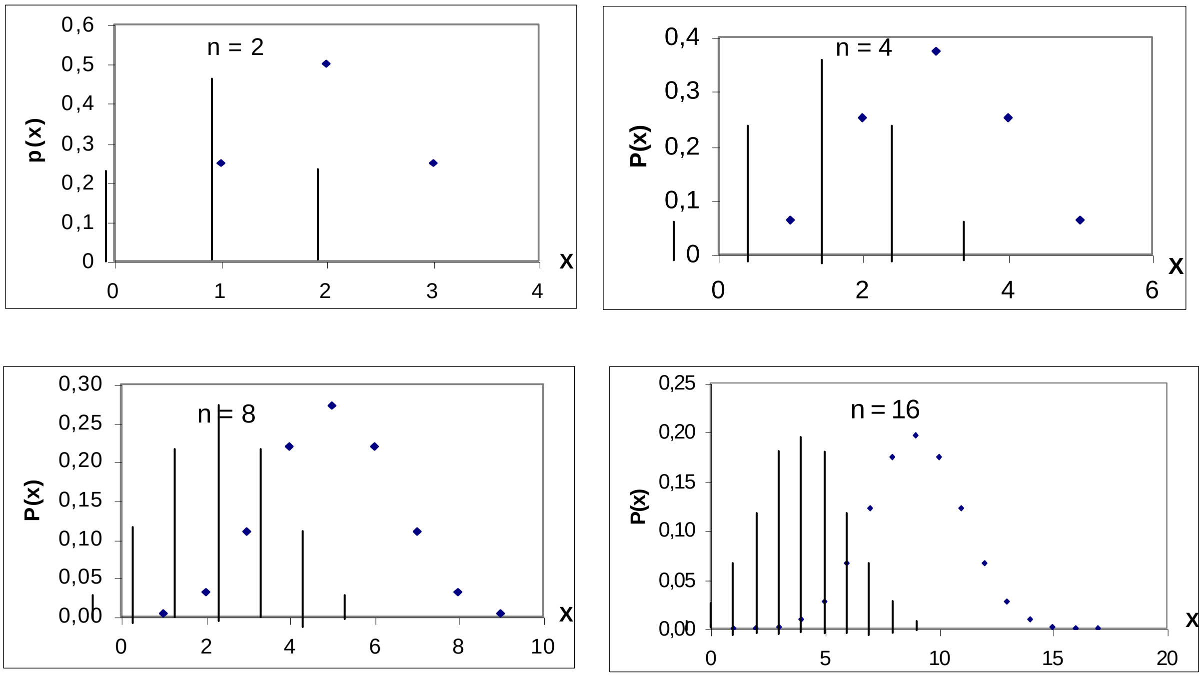
ليكن X يمثل عدد مرات الحصول على صورة عند رمي قطعة نقدية : مرتين، ٤ مرات، ٨ مرات، ١٦ مرات.

Xi	0	1	2
P_i	1/4	1/2	¼

Xi	0	1	2	3	4
P_i	1/16	4/16	6/4	4/16	1/16

Xi	0	1	2	3	4	5	6	7	8
P_i	0,004	0,031	0,109	0,219	0,273	0,219	0,109	0,031	0,004

برسم منحنيات Pi للحالات $n = 2$ ، $n = 4$ ، $n = 8$ ، $n = 16$ يظهر السلوك التقاربي للمتغير X .

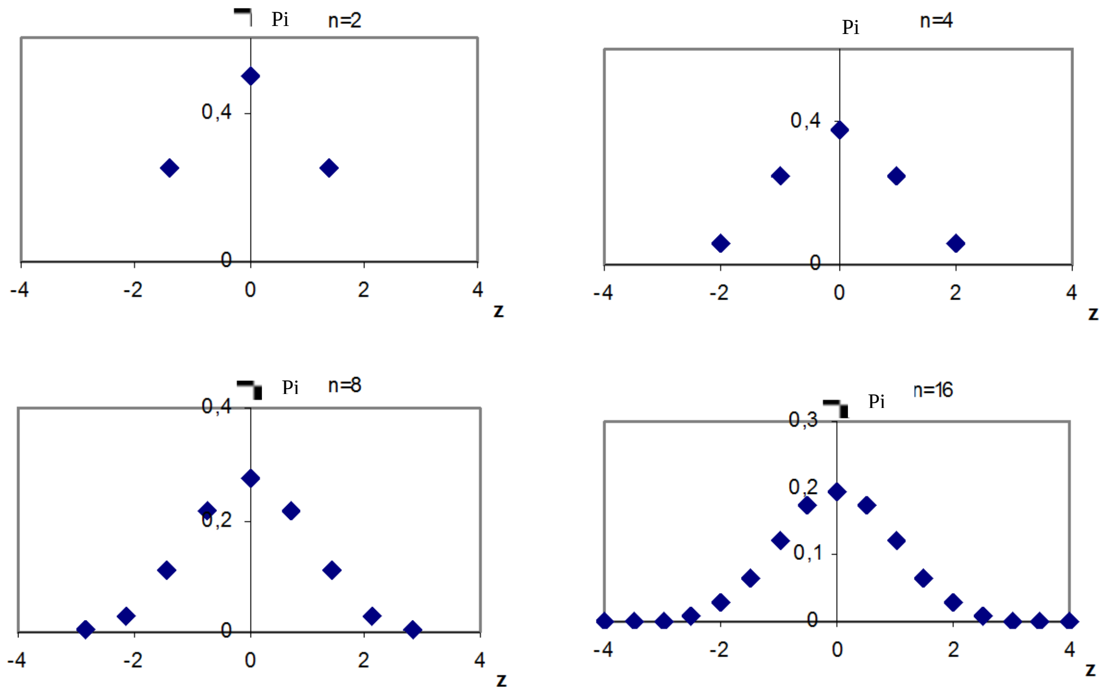


رسم 6 السلوك التقاربي للتوزيع الثنائي لما $p = 0.5$

يظهر من مقارنة المنحنيات الأربعة أن زيادة قيمة n تؤدي إلى الحصول على منحنى ذا شكل جرسي ومتماثل حول التوقع μ . هذه الملاحظة تصدق أيضا في حالة $p \neq 0.5$ لكن التحول يكون أكثر بطأ. من أجل التعميم نعتبر المتغير المعيارية $z = (x - \mu)/\sigma$ الملحقة بذات المتغير ذات التوزيع الثنائي X . إن السلوك التقاربي لـZ الملاحظ في الشكل أسفله هو ما تثبته النظرية التالية:

$$soit\ \ X \sim \mathbf{B}(n,p): \qquad Y = \frac{X - np}{\sqrt{npq}} \xrightarrow[n \to \infty]{} N(0,1).$$

ونكتب $Y \approx N(0,1)$.



رسم 7   السلوك التقاربي للتوزيع الثنائي من خلال المتغير المعيارية

قاعدة:

في حالة n كبيرة و p غير قريب من ٠ يمكن اعتبار التوزيع الثنائي كتقريب جيد للتوزيع الطبيعي. ويعطي التوزيعان نتائج أكثر تقاربا كلما كانت n كبيرة أكثر. ونكتب:

$$\lim_{n \longrightarrow \infty} P(a \le \frac{x - np}{\sqrt{npq}} \le b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{\frac{-u^2}{2}} du$$

و مما يسرع تقارب التوزيع الثنائي من التوزيع الطبيعي كون p قريب من ٠.٥ وكقاعدة :

- **عموما نعتبر أن التقريب ملائم عندما np و nq كلاهما أكبر من ٥.**
- عدد من الاحصائيين يعتمد قاعدة أخرى هي أن يكون أحد الشرطين التاليين متوفرين:
  - $npq \ge 9$
  - $n \ge 20$, $np \ge 10$, $nq \ge 10$

في حالة $p = 0.5$ ، الشرط (١) يتحقق عند $n = 36$ والثاني عند $n = 20$.

في حالة $p = 0.10$ ، الشرطين يتحققان عند $n = 100$.

الانتقال من المتغير المتقطع إلى المتغير المستمر.

لاستخدام التوزيع الطبيعي بدلا من التوزيع الثنائي يعني حساب الاحتمال عن طريق توزيع مستمر بينما المتغير متقطع. من أجل ذلك يتم اعتبار كل قيمة في المتغير الأصلي فترة.

مثال. احتمال ٤ نجاحات خلال n تجربة يصاغ كما يلي: $P(3.5 \le X \le 4.5)$.

مثال٢: نرمي قطعة نقدية ٢٠ مرة. ليكن X عدد مرات الحصول على صورة. أحسب $P(X = 8)$ ثم أدرس إمكانية استخدام نظرية موافر – لابلاس لحساب نفس الاحتمال.

$P(X = 8) = F(8) – F(7) = 0.2517 – 0.1316 = 0.1201.$, $X \sim B(20, 0.5)$

لدينا np = 10 >5 وكذلك nq = 10 >5 ، وإذا شئنا استخدام القاعدة الثانية فإننا نجد أيضا أن : n =10 ،np =10، nq=10، يمكن إذا اعتبار Y = (X-10)/√5 ~ N(0 ,1) . نستخدم المتغير المستمر X* بدلا من X لحساب احتمال الفترة المعبرة عن القيمة ٨ وهو [7.5, 8.5]

$$P(7.5 \le X^* \le 8.5) = P\left(\frac{7.5-10}{2.24} \le Z \le \frac{8.5-10}{2.24}\right) = P(-1.12 \le Z \le -6.67) = 0.12$$

التقارب بين التوزيع الثنائي وتوزيع بواسون

يعطي توزيع بواسون نتائج قريبة من التوزيع الثنائي لما $n \ge 30$ و $np < 5$ أو $nq < 5$ و يستخدم بعض الإحصائيين كشرط لاستعمال قانون بواسون بدلا من القانون الثنائي القاعدة التالية:

$p \le 0,1$ و $n \ge 25$

مثال : ١٠ % من إنتاج آلة ما يعد تالفا، نأخذ ٣٠ وحدة من انتاج هذه الآلة عشوائيا.

أحسب احتمال أن يكون هناك وحدتان تالفتان.

$P(X = 2) = C^2_{30}\ (0,1^2)\ (0.9^{28}) = 0.22$

لدينا $n \ge 25$، $p \le 0.1$: لاستعمال توزيع بواسون نحسب أولا قيمة المعلمة (معلمة قانون بواسون)

$\lambda = \mu = np = 30 * 0,1 = 3$

$P(2) = \lambda^x * e^{-\lambda}/x! = (3^2 * e^{-3}) / 2! = 0.22$

نظرية النهاية المركزية

لتكن المتغيرات X_1، X_2، . . . . متغيرات عشوائية مستقلة لها نفس التوزيع الاحتمالي بتباين ومتوسط محددين. إذا كانت

$$S_n = X_1 + X_2 + \ldots + X_n \qquad (n = 1, 2, \ldots),$$

فإن S_n تتبع التوزيع الطبيعي عندما $n \to \infty$. وبما أن $E(S_n) = n\mu$ و $\sigma_{S_n} = \sigma\sqrt{n}$ فإننا نكتب :

$$\lim_{n \longrightarrow \infty} P\left(a \le \frac{S_n - n\mu}{\sigma\sqrt{n}} \le b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{\frac{-z^2}{2}} dz$$

في الحقيقة فإن النظرية محققة عندما تكون المتغيرات المستقلة Xi لها نفس المتوسط والتباين حتى ولو لم يكن لها بالضرورة نفس التوزيع، مع العلم أنه توجد صيغ أخرى لهذه النظرية حيث لا يشترط أن يكون للمتغيرات نفس التوزيع الاحتمالي ولا حتى أن تكون مستقلة. تجدر الإشارة إلى أن نظرية موافر- لابلاس التي تطرقنا إليها سابقا هي حالة خاصة من نظرية النهاية المركزية، ذلك أن المتغير تتبع القانون B(n, p) يمكن اعتبارها مجموعا لعدد من المتغيرات المستقلة ذات التوزيع البرنولي B(1, p).

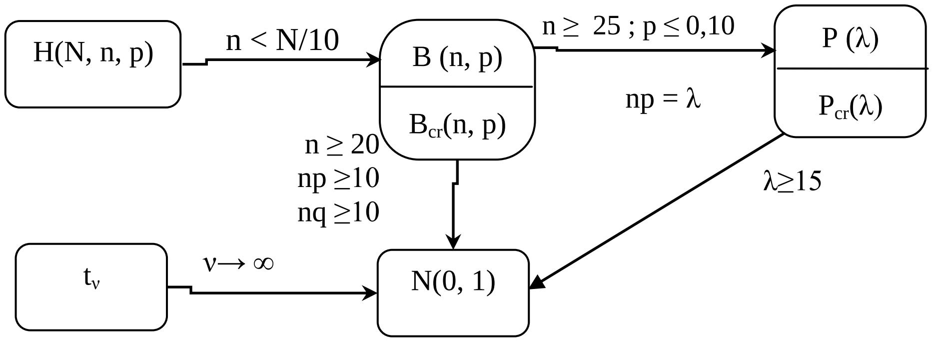
خلاصة

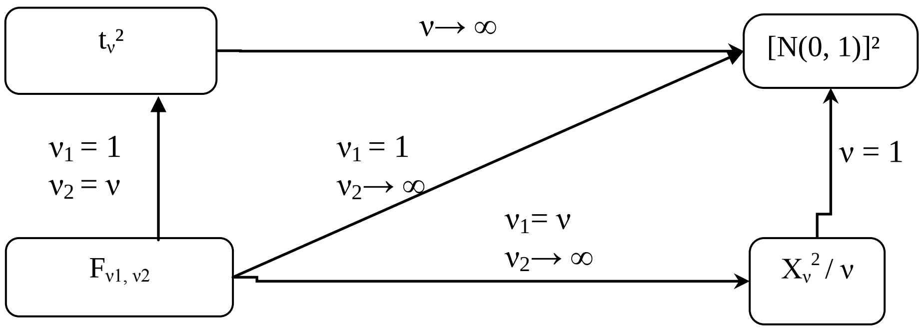
لاستخدام التوزيع الطبيعي بدلا من التوزيع الثنائي يعني حساب الاحتمال عن طريق توزيع مستمر بينما المتغير متقطع. من أجل ذلك يتم اعتبار كل قيمة في المتغير الأصلي فترة.

نظرية النهاية المركزية تنص على أن S_n (متغيرات عشوائية مستقلة لها نفس التوزيع الاحتمالي بتباين ومتوسط محددين) تتبع التوزيع الطبيعي عندما $n \to \infty$ بمتوسط $E(S_n) = n\mu$ و $\sigma_{S_n} = \sigma\sqrt{n}$ ونكتب:

$$\lim_{n \longrightarrow \infty} P\left(a \le \frac{S_n - n\mu}{\sigma\sqrt{n}} \le b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{\frac{-z^2}{2}} dz$$

الرسم البياني التالي يبين القواعد المستخدمة كشروط للتقريب بين التوزيعات الاحتمالية المذكورة آنفا في المبحث بالإضافة إلى التوزيعات الأخرى التي درست في الفصول السابقة (الرمز cr يعني المتغير معيارية).





رسم 8 يبين قواعد التقريب بين القوانين الاحتمالية

الفصل الثاني

نظرية توزيع المعاينة

تنتشر في مجتمعاتنا المعاصرة عمليات الاستقصاء، ففي عالم الأعمال تقوم المؤسسات عن طريق مصالح التسويق ومصالح البحوث والتطوير بإجراء استقصاءات للإطلاع على توجهات المستهلكين، وفي وسائل الإعلام لا يمر يوم دون أن يعلن عن نتائج استقصاء أجرته مجلة أو جامعة حول مواضيع سياسية أو اجتماعية متعددة، منها الاستقصاءات المثيرة للجدل حول الأراء السياسية للمواطنين أثناء الحملات الانتخابية. فما هي الأسس النظرية الرياضية التي تستند عليها الاستقصاءات المختلفة ؟ أو كيف يمكن الاستدلال من خلال بيانات عينة على خصائص المجتمع الذي أخذت منه؟ الإجابة على هذه الأسئلة و غيرها تتطلب فهم العلاقات الرياضية بين الخصائص المختلفة للمجتمع مثل المتوسط، التباين وغيرها، والخصائص المناظرة لها في العينة وهو ما سنتناوله في هذا الفصل. في الفصول المقبلة سندرس عددا من التطبيقات لهذه العلاقات الرياضية.

مفاهيم إحصائية

المجتمع والعينة Population and sample

نشرح هذين المصطلحين من خلال الأمثلة التالية:

- قد ترغب الإدارة العسكرية في تقدير الوزن المتوسط للجندي، فتقوم أخذ أوزان عينة من ١٠٠ جندي من بين مجموع الجنود (المجتمع).
- ترغب هيأة معينة بالبحوث السياسية في تقدير نسبة الناخبين المساندين لمرشح معين في ١٠ الولايات، فتقوم باستجواب ١٠٠ ناخب من كل ولاية. الناخبون في الولايات العشر يمثلون المجتمع بينما ال ١٠٠٠ ناخب المستجوبون يمثلون العينة.
- من أجل معرفة مدى دقة صنع قطعة نقدية ترمى القطعة ١٠٠ مرة ونحسب عدد مرات الحصول على الصورة والكتابة، حجم العينة هنا هو ١٠٠.
- لتقدير نسبة الكرات داخل صندوق، التي من لون معين، نقوم عدد من المرات بسحب كرة نسجل لونها ثم نعيدها. عدد الكرات المسحوبة يمثل حجم العينة.

نلاحظ أن مصطلح المجتمع يقصد به القياسات أو القيم وليس الأفراد أو الأشياء التي تم قياسها (مجتمع الأوزان، مجتمع آراء الناخبين..)، كما أن المجتمع قد يكون محدودا أو غير محدود (نتائج رميات قطعة النقد)، أما العينة فهي عادة تكون محدودة، ونرمز عادة لحجم المجتمع ب N، ولحجم العينة ب n.

العينة النفادية والعينة غير النفادية

عندما يكون السحب بالإرجاع حيث يمكن أن تظهر المفردة أكثر من مرة في العينة، نسمي هذه المعاينة غير نفادية لأن تكرار العملية لا يؤدي إلى تقليص عدد المفردات في المجتمع، والعكس نسمي المعاينة بدون إرجاع معاينة نفادية. هناك فرضيتان تتكرران في عدد من العلاقات الرياضية التي سنراها لاحقا، هما فرضية أن قيم مفردات العينة مستقلة والمجتمع لانهائي. يتحقق شرط الاستقلال إذا كانت المعاينة غير نفادية، وإذا كانت كذلك، يمكن اعتبار المجتمع مجتمعا غير محدود.

العينة العشوائي Random Sample

من أجل أن تكون العينة ممثلة للمجتمع، أحد الطرق المستخدمة هي العينة العشوائي. نظريا (قد يصعب تحقيق ذلك في الواقع)، نقول عن عينة أنها عشوائية إذاكان لكل مفردة في المجتمع نفس الاحتمال لأن تكون في العينة. تسمى هذه العينة بالعينة العشوائي البسيطة. لإنجاز ذلك إما أن نسحب المفردات بطريقة عشوائية أو نرقم مفردات المجتمع ثم نحدد العينة من خلال مجموعة من الأعداد تؤخذ من الجداول الإحصائية للأعداد العشوائي.

معالم المجتمع Parametrical population

نقصد بمعالم المجتمع مجموعة من خصائصه مثل المتوسط، التباين، معامل التماثل، ... من خصائص المجتمع أيضا طبيعة توزيعه الاحتمالي $f(x)$ كأن يكون طبيعيا أو غيره.

إحصائية المعاينة Sampling Statistic

لتقدير معالم المجتمع (متوسط المجتمع μ، تباين المجتمع σ^2 ، النسبة p ...) ننطلق من بيانات العينة، حيث نحتاج إلى حساب معالم مثل متوسط العينة m ، تباين العينة S^2، النسبة في العينة p'. بصفة عامة، نسمي كل قيمة تحسب انطلاقا من بيانات العينة من أجل تقدير قيمة معالم المجتمع إحصائية المعاينة. نظريا (رياضيا) إحصائية المعاينة هي كل دالة في المتغيرات العشوائي التي تمثل القيم المحصل عليها في العينة.

توزيع المعاينة للمتوسطات

متوسط توزيع المعاينة للمتوسطات

مسألة: ليكن المجتمع 1، 3، 5، 6، 8. ما هي القيمة المتوقعة لمتوسط عينة مسحوبة بالإرجاع مكونة من مفردتين (m)؟ أحسب متوسط المجتمع μ. قارن بين m و μ. من أجل تحديد ذلك أحسب جميع الحالات الممكنة للمتوسط m_i حسب كل عينة.

العينات الممكنة العينات الممكنة ذات الحجم $n = 2$ من مجتمع حجمه ٥ عددها: ٥×٥ =٢٥

المتوسطات الممكنة للعينة (معاينة غير نفادية) m_i				
1	2	3	3,5	4,5
2	3	4	4,5	5,5
3	4	5	5,5	6,5
3,5	4,5	5,5	6	7
4,5	5,5	6,5	7	8

العينات الممكنة				
(1, 1)	(3, 1)	(5, 1)	(6, 1)	(8, 1)
(1, 3)	(3, 3)	(5, 3)	(6, 3)	(8, 3)
(1, 5)	(3, 5)	(5, 5)	(6, 5)	(8, 5)
(1, 6)	(3, 6)	(5, 6)	(6, 6)	(8, 6)
(1, 8)	(3, 8)	(5, 8)	(6, 8)	(8, 8)

القيمة المتوقعة m ل m_i هي متوسط قيمها وهي $m = (\sum_i m_i) / 25 = 4,6$.

حساب متوسط المجتمع: $\mu = (1 + 3 + 5 + 6 + 8)/5 = 4.6$

مثال٢. أوجد نفس مطالب المثال ١. في حالة السحب بدون إرجاع. العينات الممكنة عددها: $C^2_5 = 10$

المتوسطات الممكنة للعينة أو **توزيع المعاينة للمتوسطات (معاينة نفادية)** m_i			
2			
3	4		
3,5	4,5	5,5	
4,5	5,5	6,5	7

العينات الممكنة بدون إرجاع			
(1, 3)			
(1, 5)	(3, 5)		
(1, 6)	(3, 6)	(5, 6)	
(1, 8)	(3, 5)	(5, 8)	(6, 8)

القيمة المتوقعة m ل m_i هي متوسط قيمها وهي:

$E(m) = \mu_m = (\sum_i m_i) / 10 = 4,6$

متوسط المجتمع : $\mu = (1 + 3 + 5 + 6 + 8)/5 = 4.6$

نظرية ١. إذا كانت م ع تمثل مجتمع ما وm المتغير ع تمثل متوسط عينة مسحوبة من ذات المجتمع، فإن القيمة المتوقعة لمتوسط العينة E(M) تكتب كما يلي: $E(M) = \mu_m = \mu$

البرهان : لنرمز ب Xi لقيم المتغير الأصلية X .

$$E(\overline{X}) = E\left(\frac{1}{n}\sum_i Xi\right) = \frac{1}{n}\sum_i E(Xi) = \frac{1}{n}\sum_i \mu = \frac{1}{n}n\mu = \mu.$$

تباين توزيع المعاينة للمتوسطات

حالة المعاينة بالإرجاع

مثال. أحسب تباين المجتمع في المسألة ١، أحسب التباين (والانحراف المعياري) لتوزيع المعاينة للمتوسطات σ^2_m علما أن العينة مسحوبة بالإرجاع (غ نفادية)، قارن بين تباين المجتمع وتباين متوسطات العينات الممكنة (توزيع المعاينة للمتوسطات).

mi				
1	2	3	3,5	4,5
2	3	4	4,5	5,5
3	4	5	5,5	6,5
3,5	4,5	5,5	6	7
4,5	5,5	6,5	7	8

$\sigma^2_m = [\sum_i (m_i - m)^2]/25 = 2.92;$

$\sigma^2 = [\sum_i (x_i - \mu)^2]/5 = 5.84$

2.92 = 5.84 / 2

هذا المثال يمهد للنظرية التالية:

نظرية ٢. إذا كانت م ع تمثل مجتمع ما وm_i المتغير ع تمثل متوسط عينة مسحوبة من ذات المجتمع بالإرجاع، فإن تباين m_i (تباين توزيع المعاينة للمتوسطات) يكتب كما يلي: $\sigma_m^2 = \dfrac{\sigma^2}{n}$ حيث n حجم العينة.

البرهان: لنرمز ب Xi لقيم المتغير الأصلية X .

$$V(\overline{X}) = V\left(\frac{1}{n}\sum_i Xi\right) = \frac{1}{n^2}\sum_i V(Xi) = \frac{1}{n^2}\sum_i \sigma^2 = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}.$$

حالة المعاينة بدون إرجاع.

مسألة: في المسألة ١ أحسب تباين المتوسطات الممكنة للعينة σ^2_m في حالة المعاينة بدون إرجاع، قارن بين تباين المجتمع وتباين المتوسطات الممكنة للعينة.

المتوسطات الممكنة للعينة أو توزيع المعاينة للمتوسطات (معاينة نفادية) Mi			
2			
3	4		
3,5	4,5	5,5	
4,5	5,5	6,5	7

تباين المتوسطات الممكنة للعينة:

$$\sigma^2_m = [\sum_i (m_i - m)^2]/10 = 2.19$$

تباين المجتمع: $\sigma^2 = [\sum_i (x_i - \mu)^2]/5 = 5.84$

(أو بطريقة ثانية:

$$\sigma^2 = E(X^2) - E(X)^2$$
$$= (1 + 9 + 25 + 36 + 64) / 5 - 4.6^2 = 5.84)$$

المقارنة بين تباين متوسط العينة و تباين المجتمع:

$$2.19 = \frac{5.84}{2}\left(\frac{5-2}{5-1}\right)$$

هذا يمهد للنظرية التالية:

نظرية ٣. إذا كانت X م ع تمثل مجتمع ما حجمه N وm_i المتغير ع تمثل متوسط عينة حجمها n مسحوبة من ذات المجتمع بدون إرجاع، فإن تباين m_i (تباين توزيع المعاينة للمتوسطات) يكتب كما يلي: $\sigma_m^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$

وتسمى النسبة $\frac{N-n}{N-1}$ معامل الإرجاع.

طبيعة توزيع m

ندرس طبيعة توزيع متوسط توزيع المعاينة للمتوسطات من خلال النظريات التالية:

نظرية ٤. إذا كان المجتمع موزع طبيعيا بمتوسط μ وتباين σ^2 فإن متوسط العينة المسحوبة منه يتبع أيضا التوزيع الطبيعي بمتوسط μ وتباين σ^2/n، ونكتب $m \approx N(\mu, \sigma^2/n)$

نظرية ٥. (**نظرية النهاية المركزية**): إذا كان المجتمع الذي تسحب منه العينة ذو متوسط μ وتباين σ^2 لكن ليس بالضرورة طبيعيا فإن المتغير المعيارية ل m أي $z = \dfrac{m-\mu}{\sigma/\sqrt{n}}$ تؤول إلى التوزيع الطبيعي المعياري عندما يكون n كبيرا $(n \geq 30)$ ونكتب:

$.z \approx N(0, 1)$

في حالة المجتمع محدود والمعاينة نفادية نستبدل العبارة $\sigma/\sqrt{n}$ ب $\sigma_m = \dfrac{\sigma}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}}$

عمليا يستخدم الإحصائيين هذه الصيغة المعدلة بمعامل الإرجاع للانحراف المعياري عندما $n/N \geq 0.05$

مثال: مجتمع حجمه 900 بمتوسط μ =٢٠ و12= σ . نستخرج كل العينات الممكنة. أحسب المتوسط والانحراف المعياري لتوزيع المعاينة للمتوسطات في حالة: (١) حجم العينة n = 36، (٢) n = 64 .

(1) $n = 36$: $n/N = 36/900 = 0.04 < 0.05 \Rightarrow \sigma_m = \sigma/\sqrt{n} = 12/\sqrt{36} = 2$

(2) $n = 64 : N = 900 \Rightarrow \dfrac{n}{N} = \dfrac{64}{900} = 0.071 > 0.05$

$$\Rightarrow \sigma_m = \frac{12}{\sqrt{64}}\sqrt{\frac{900-64}{900-1}} = 1.92$$

$E(m) = \mu = 20$

مثال٢. باستخدام معطيات المثال السابق (n = 36) أحسب احتمال أن يكون m محصورا بين ١٨ و٢٢.

أحسب نفس الاحتمال في حالة n = 64.

$$Z_1 = \frac{m_1 - \mu}{\sigma/\sqrt{n}} = \frac{18-20}{12/\sqrt{36}} = -1\ ,\quad Z_2 = 1 \quad \Rightarrow P(18 < m < 22) = P(Z_1 < Z < Z_2) = 0.6827$$

$$Z_1 = \frac{m_1 - \mu}{\sigma/\sqrt{n}\left(\dfrac{N-n}{N-1}\right)} = \frac{18-20}{1.92} = -1.04\ ,\quad Z_2 = 1.04 \quad \Rightarrow P(18 < m < 22) = P(-1.04 < Z < 1.04) = 0.70$$

خلاصة :الجدول التالي يبين أهم خصائص توزيع المعينة للمتوسطات.

المجتمع	المعاينة	الخاصية
مجتمع ما	سحب بالإرجاع أو بدون إرجاع	$E(M) = \mu_m = \mu$
مجتمع ما	سحب بالإرجاع	$\sigma_m^2 = \frac{\sigma^2}{n}$
مجتمع ما حجمه N	سحب بدون إرجاع	$\sigma_m^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$
مجتمع موزع طبيعيا بمتوسط μ وتباين σ^2	سحب بالإرجاع أو بدون إرجاع	$m \approx N(\mu, \sigma^2/n)$
مجتمع بمتوسط μ وتباين σ^2 لكن ليس بالضرورة طبيعيا	عندما يكون n كبيرا $(n \geq 30)$	$z = \frac{m-\mu}{\sigma/\sqrt{n}} \approx N(0, 1)$

توزيع المعاينة للنسبة

النظرية التالية تبين المتوسط، التباين، و طبيعة التوزيع الإحصائية **p'** : نسبة خاصية ما في العينة.

نظرية : ٦ لتكن X م ع تمثل مجتمع ما غير محدود وموزع طبيعيا حيث p نسبة المفردات في المجتمع ذات صفة معينة، ولتكن p' تمثل نسبة المفردات ذات الصفة المذكورة في العينة المسحوبة من ذات المجتمع، نحصل على توزيع للإحصائية p' حيث معالمه $E(p')$ و$u_{p'}$، هذه المعالم تساوي :

$$E(p') = \mu_{p'} = p \qquad ; \qquad \sigma^2_{p'} = \frac{pq}{n}$$

عند $n \geq 30$: $p' \approx N(p, \sigma_{p'})$

عندما يكون المجتمع محدودا والمعاينة نفاديه نضرب في معامل الإرجاع عند حساب الانحراف المعياري.

مثال. لاحظت إدارة الجامعة أنه في عينة من ١٠٠ طالب، ٤٠ حصلوا أخيرا على شهادة. تريد الإدارة تقدير نسبة الطلبة الذين يحصلون على الشهادة داخل فترة يكون احتماله ٩٠ بالمائة.

$P(p_1 < p' < p_2) = 0.9$; $n \geq 30$,

نفترض أن N كبير بحيث : $n/N < 0.05$

$$\Rightarrow p' \sim N(p, \sigma_{p'}),\ \sigma_{p'} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.4(0.6)}{100}} \cong 0.05$$

$$P(p_1 < p' < p_2) = 0.9 = P(z_1 < Z < z_2) \Rightarrow z_1 = -1.64, \qquad z_2 = 1.64$$

$$Z1 = \frac{(p_1 - p)}{\sigma_{p'}} \Rightarrow p = p' \pm z(\sigma_{p'}) = 0.4 \pm 1.64(0.05) = 0.4 \pm 0.082$$

$$\Rightarrow P(0.318 < p < 0.482) = 0.9.$$

توزيع المعاينة للفروق والمجاميع

المتوسط والتباين

ليكن لدينا مجتمعين نسحب من كل منهما عينة عشوائية، نحسب في كل عينة محسوبة من المجتمع الأول الإحصائية S_1 ونحسب نفس الاحصائية (المتوسط مثلا أو التباين ...) في كل عينة من المجتمع الثاني ونسميها S_2. إن الفرق $S_1 - S_2$ يشكل بدوره المتغير عشوائية لها المتوسط والتباين التاليين:

$$\mu_{S-S2} = \mu_{S1} - \mu_{S2} \qquad\qquad \sigma^2_{S1-S2} = \sigma^2_{S1} + \sigma^2_{S2}$$

مثال ١. إذا كانت الاحصائية هي **المتوسط** فإن:

$$\mu_{m1-m2} = \mu_{m1} - \mu_{m2} = \mu_1 - \mu_2 \qquad\qquad \sigma^2_{m1-m2} = \sigma^2_{m1} + \sigma^2_{m2} = \sigma^2/n_1 + \sigma^2/n_2$$

مثال ٢. إذا كانت الاحصائية هي **النسبة** فإن:

$$\mu_{p1-p2} = \mu_{p1} - \mu_{p2} = p_1 - p_2 \qquad\qquad \sigma^2_{p1-p2} = \sigma^2_{p1} + \sigma^2_{p2} = p_1q_1/n_1 + p_2q_2 / n_2$$

إذا كان الاهتمام هو على مجموع الاحصائيتين بدلا من الفرق بينهما فإن:

$$\mu_{S1+S2} = \mu_{S1} + \mu_{S2} \qquad\qquad \sigma^2_{S1+S2} = \sigma^2_{S1} + \sigma^2_{S2}$$

طبيعة توزيع المعاينة للفرق بين متوسطين

نظرية ٧ **:** في حالة $n_1 \geq 30$ وn_2 ، يقترب توزيع المتغير المعيارية للفرق بين متوسطين من التوزيع الطبيعي المعياري.

ونكتب: $\mu_{m1-m2} \approx N(0, 1)$

مثال : ليكن المجتمع U_1 : ٣، ٧، ٨. والمجتمع U_2 : ٢، ٤. تحقق من أن :

$$\mu_{U1-U2} = \mu_{U1} - \mu_{U2} ; \qquad \sigma^2_{U1-U2} = \sigma^2_{U1} + \sigma^2_{U2} .$$

$\mu_{U1} = (3 + 7 + 8)/3 = 6$; $\mu_{U2} = (2 + 4)/2 = 3 \Rightarrow$

$\mu_{U1} - \mu_{U2} = 6 - 3 = 3$

$\mu_{U1-U2} = (1 + 5 + 6 - 1 + 3 + 4)/6 = 3$

$\sigma^2_{U1} = (3^2 + 7^2 + 8^2)/3 - 6^2 = 14/3$;

$\sigma^2_{U2} = (2^2 + 4^2)/2 - 3^2 = 1 \Rightarrow \sigma^2_{U1} + \sigma^2_{U2} = 17/3$

$\sigma^2_{U1-U2} = (1^2 + 5^2 + 6^2 + 1^2 + 3^2 + 4^2) / 6 - 3^2 =$

$(1 + 25 + 36 + 1 + 9 + 16) / 6 - 9 = 17/3$

U_1				
8	7	3	**$U_1 - U_2$**	
6	**5**	**1**	2	U_2
4	**3**	**-1**	4	

توزيع المعاينة للتباين وتوزيع المعاينة لنسبة تبايني عينتين

توزيع المعاينة للتباين

حالة المعاينة بالإرجاع

مسألة: أحسب تباين المجتمع في المسألة ١، أحسب القيمة المتوقعة لتباين العينة المسحوبة بالإرجاع من خلال متوسط تباينات العينات الممكنة، قارن بين تباين المجتمع والقيمة المتوقعة لتباين العينة.

التباينات الممكنة S^2_i				
0	1	4	6,25	12,3
1	0	1	2,25	6,25
4	1	0	0,25	2,25
6,25	2,25	0,25	0	1
12,3	6,25	2,25	1	0

$(\sum_i S^2_i)/25 = 73/25 = 2.92 \quad \Rightarrow E(S^2) = 2.92$

$\sigma^2 = E(X^2) - E(X)^2$

$= [(1 + 9 + 25 + 36 + 64)/5] - 4.6^2 = (135/5) - 21 = 5.84$

$\sigma^2 (1/n) \quad = \quad E(S^2) = 2.92 = 5.84/2$

نظرية ٨ : إذا كانت م ع تمثل مجتمع ما وS^2 المتغير ع تمثل تباين عينة مسحوبة بالإرجاع (أو بدون إرجاع من مجتمع غير محدود) حجمها n، فإن :

$$E(S^2) = \mu_{S^2} = \sigma^2\left(\frac{n-1}{n}\right)$$

($E(S^2) \approx \sigma^2$: $n \geq 30$ عند)

البرهان:

$$E(S^2) = E\left[\frac{1}{n}\sum_i (x_i - \bar{x})^2\right] = E\left[\frac{1}{n}\sum_i x_i^2 - \bar{x}^2\right] = \frac{1}{n}\sum_i E(x_i^2) - E(\bar{x}^2)$$

$$= \frac{1}{n}\sum_i \left(V(X) + \mu^2\right) - [V(\bar{x}) + E(\bar{x})^2] = \sigma^2 + \mu^2 - \frac{\sigma^2}{n} - \mu^2 = \sigma^2(1 - \frac{1}{n}) = \sigma^2\left(\frac{n-1}{n}\right)$$

ملاحظة: من النظرية نجد أن: $E\left(S^2\frac{n}{n-1}\right) = \sigma^2$ ونقول عن $S^2\frac{n}{n-1}$ أنه مقدر "غير منحرف" ل σ^2 **و**يرمز له ب $S^{2\prime}$

حيث:

$$\hat{S}^2 = S^2\frac{n}{n-1}$$

نظرية ٩ : إذا أخذنا عينات عشوائية حجمها n من مجتمع طبيعي، فإن : $\frac{nS^2}{\sigma^2} = \frac{(n-1)\hat{S}^2}{\sigma^2} \sim \chi^2_{n-1}$

مثال : ليكن مجتمع طبيعي حجمه 100 نسحب منه عينة حجمها n = 16 . ما هو احتمال أن يكون تباين العينة S^2 أقل من أو يساوي ١٠ علما أن تباين المجتمع ٨٠.

$$P(S^2 \le 10) = P(\chi^2_{15} \le 2) = P(S^2 \frac{16}{80} \le X \sim N(\mu,\sigma)) \Rightarrow \frac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$$

$P(X^2_{15} \le 2) < 0.005$ من الجدول

حالة المعاينة بدون إرجاع

مسألة: في المسألة ١ أحسب تباين المتوسطات الممكنة للعينة σ^2_m في حالة المعاينة بدون إرجاع، قارن بين تباين المجتمع وتباين المتوسطات الممكنة للعينة.

التباينات الممكنة S^2_i			
1			
4	1		
6,25	2,25	0,25	
12,3	6,25	2,25	1

$(\sum_i S^2_i) = ٣٦.٥$; $(\sum_i S^2_i)/١٠ = ٣.٦٥$ => $E(S^2) =$ **٣.٦٥**

$\sigma^2 = E(X^2) - E(X)^2$

$= [(1 + 9 + 25 + 36 + 64)/5] - 4.6^2 = 5.84$

$E(S^2) =$ ٣.٦٥ = ٥.٨4*(5/٤) (1/2)

$= \sigma^2 * [(n-1)/ n] [N/ (N-1)]$

نظرية ١٠ : إذا كانت م ع تمثل مجتمع ما محدود وS^2 المتغير ع تمثل تباين عينة نفادية مسحوبة من ذات المجتمع، فإن القيمة المتوقعة لتباين العينة تكتب: $E(S^2) = \mu_{S^2} = \sigma^2\left(\frac{n-1}{n}\right)\left(\frac{N}{N-1}\right)$

(عندما يكون N كبير جدا N/ (N-1) تؤول إلى ١)

الانحراف المعياري لتوزيع المعاينة للتباين

نظرية ١١ : إذا كانت X م ع تمثل مجتمع ما وS^2 المتغير ع تمثل تباين عينة مسحوبة من ذات المجتمع، فإن:

$$\sigma_{S^2} = \begin{cases} \sigma^2\sqrt{2/n} & \text{For } X\sim N \\ \sqrt{\dfrac{\mu^4 - \sigma^4}{n}} & \text{otherwise} \end{cases}$$

من أجل $n \ge 100$، توزيع S^2 يقترب كثيرا من التوزيع الطبيعي.

(د) الانحراف المعياري لتوزيع المعاينة للانحراف المعياري

$$\sigma_S = \begin{cases} \dfrac{\sigma}{\sqrt{2/n}} & \text{For } X \sim N \\ \sqrt{\dfrac{\mu^4 - \sigma^4}{4n\sigma^2}} & \text{otherwise} \end{cases}$$

من أجل $n \geq 100$، توزيع S يقترب كثيرا من التوزيع الطبيعي و$S \approx \mu_S$.

توزيع المعاينة لنسبة تباينين

رأينا في الفصل السابق أن: $X = \dfrac{X_1/\nu_1}{X_2/\nu_2} \sim F_{\nu_1, \nu_2}$ في حالة المتغيران العشوائيتان مستقلان و$X_1 \sim \chi^2_{\nu_1}$ و $X_2 \sim \chi^2_{\nu_2}$. من النظرية ٩ نستنتج ما يلي:

نظرية ١٢ : ليكن لدينا مجتمعان طبيعيان تبايناهما σ^2_1 , σ^2_2 . نسحب من المجتمعين عينتين عشوائيتين حجمهما على التوالي n_1 , n_2 :

$$F = \frac{\left[\dfrac{S_1^2 n_1}{n_1-1}\right]\dfrac{1}{\sigma_1^2}}{\left[\dfrac{S_2^2 n_2}{n_2-1}\right]\dfrac{1}{\sigma_2^2}} = \frac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2} \to F_{n_1-1;\, n_2-1}$$

مثال . عينتين حجمهما ٨ و١٠ مسحوبتين من مجتمعين طبيعيين تبايناهما على التوالي ٢٠ و٣٦. ما احتمال أن يكون تباين الأولى أكبر من ضعف تباين الثانية؟

$$P(S_1^2 > 2S_2^2) = P(\frac{S_1^2}{S_2^2} > 2) = P\left(\frac{S_1^2\left(\dfrac{n_1}{n_1-1}\right)\dfrac{1}{\sigma_1^2}}{S_2^2\left(\dfrac{n_2}{n_2-1}\right)\dfrac{1}{\sigma_2^2}} > 2\frac{\left(\dfrac{n_1}{n_1-1}\right)\dfrac{1}{\sigma_1^2}}{\left(\dfrac{n_2}{n_2-1}\right)\dfrac{1}{\sigma_2^2}}\right) = P\left(\frac{S_1^2\left(\dfrac{n_1}{n_1-1}\right)\dfrac{1}{\sigma_1^2}}{S_2^2\left(\dfrac{n_2}{n_2-1}\right)\dfrac{1}{\sigma_2^2}} > 2\frac{\left(\dfrac{8}{7}\right)\dfrac{1}{20}}{\left(\dfrac{10}{9}\right)\dfrac{1}{36}}\right) =$$

$$= P(F_{7,9} > 3.7)$$

من الجدول نجد $0.05 > P(F_{7,9} > 3.7) > 0.01$ و في الحقيقة $P(F_{7,9} > 3.7) = 0.036$

خلاصة

الجدول التالي يلخص ما ورد في النظريات السابقة من ٦ إلى ١٠ .

إحصائية العينة	المجتمع	المعاينة	الخاصية
النسبة	مجتمع موزع طبيعيا غير محدود		$E(p')=\mu_{p'}=p$; $\sigma^2{}_{p'}=\dfrac{pq}{n}$
		$n \geq 30$	$p' \approx N(p, \sigma_{p'})$
	مجتمع طبيعي محدود	المعاينة نفاديه	لحساب $\sigma_{p'}$ نضرب في معامل الإرجاع.
الفرق بين إحصائيتين ما.	مجتمع ما	سحب بالإرجاع	$\mu_{S-S2}=\mu_{S1}-\mu_{S2}$ $\mu_{S1+S2}=\mu_{S1}+\mu_{S2}$ $\sigma^2_{S1-S2}=\sigma^2_{S1}+\sigma^2_{S2}$ $\sigma^2_{S1+S2}=\sigma^2_{S1}+\sigma^2_{S2}$
		n_1 و $n_2 \geq 30$	$\mu_{m1-m2} \approx N(0, 1)$
التباين	مجتمع ما وتباين عينة S^2	سحب بالإرجاع (أو بدون إرجاع من مجتمع غير محدود) حجمها n	$E(S^2)=\mu_{S^2}=\sigma^2\left(\dfrac{n-1}{n}\right)$
		$n \geq 30$	$E(S^2) \approx \sigma^2$
	مجتمع طبيعي	حجمها n	$\dfrac{nS^2}{\sigma^2}=\dfrac{(n-1)\hat{S}^2}{\sigma^2}\sim\chi^2_{n-1}$
	مجتمع ما محدود وS^2 تمثل تباين العينة	<u>عينة نفادية</u>	$E(S^2)=\mu_{S^2}=\sigma^2\left(\dfrac{n-1}{n}\right)\left(\dfrac{N}{N-1}\right)$
		N كبير جدا	N/ (N-1) تؤول إلى ١
نسبة تباينين	مجتمعان طبيعيان تبايناهما σ^21 , σ^22	عينتين عشوائيتين حجمهما على التوالي n_1 , n_2	$F=\dfrac{\left[\dfrac{S_1^2 n_1}{n_1-1}\right]\dfrac{1}{\sigma_1^2}}{\left[\dfrac{S_2^2 n_2}{n_2-1}\right]\dfrac{1}{\sigma_2^2}}=\dfrac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2}\to F_{n_1-1;\, n_2-1}$

الفصل الثالث

نظرية التقدير

في الفصل السابق درسنا من خلال مجموعة من النظريات العلاقة الرياضية بين معالم العينة والمعالم المناظرة لها في المجتمع مثل المتوسط، التباين، النسبة...كما درسنا العلاقة بين شكل توزيع المجتمع وشكل التوزيع الاحتمالي لمعالم العينة. تظهر هذه العلاقات كتوصيف لخصائص العينة ومعالمها ولكنها تستخدم أكثر لتقدير خصائص ومعالم المجتمع محل الدراسة، وهذا ما سنتعرف عليه في هذا الفصل.

بعض خصائص المقدر

لتقدير معلمة من معالم مجتمع محل دراسة، نحتاج إلى اختيار الإحصائية المناسبة في العينة لتقدير هذه المعلمة. غالبا ما تكون المعلمة المناظرة في العينة هي أحسن مقدر، كأن نقدر متوسط المجتمع μ من خلال متوسط العينة μ_m. تسمى الإحصائية المستخدمة في التقدير المقدر.

المقدر غير المتحيز Unbiased

نقول عن إحصائية ما بأنها مقدر غير متحيز لمعلمة المجتمع إذا كان متوسطها أو توقعها الرياضي مساويا لمعلمة المجتمع.

مثال: نقول عن متوسط العينة m أنه مقدر غير متحيز لمتوسط المجتمع μ لأن $E(m) = \mu$. في المقابل نسمي الإحصائية S^2 في معاينة بالإرجاع أنها مقدر متحيز ل σ^2 لأن $E(S^2) = \sigma^2 (n-1)/n \neq \sigma^2$ ، بينما تعتبر الاحصائية $S'^2 = S^2 n/(n-1)$ مقدرا غير متحيز في معاينة بالإرجاع.

الكفاءة Efficient

تتعلق كفاءة مقدر ما بمقدار التباين لتوزيع المعاينة للإحصائية، فإذا كان لمقدرين (إحصائيتين) نفس المتوسط نقول عن المقدر ذو توزيع المعاينة الأقل تباينا أنه الأكثر كفاءة.

مثال: لكل من توزيعي المعاينة للمتوسط والوسيط نفس المتوسط هو متوسط المجتمع μ ، لكن يعتبر المتوسط m مقدرا أكثر كفاءة لمتوسط المجتمع μ من الوسيط لأن تباين توزيع المعاينة للمتوسطات $V(m) = \sigma^2/n$ أقل من تباين توزيع المعاينة للوسيط :

$$V(méd) = \sigma^2\pi/2n = (\sigma^2/n)\,(3.14159/2) \quad > \quad \sigma^2/n.$$

من البديهي أن استخدام مقدرات فعالة وغير متحيزة هو الأفضل، إلا أنه قد يلجأ لمقدرات أخرى لسهولة الحصول عليها.

التقارب Convergence

نقول عن مقدر أنه متقارب إذا كان يؤول إلى قيمة المعلمة المقدرة عندما يؤول حجم العينة إلى ما لا نهاية.

مثال: يعتبر متوسط العينة مقدرا متقاربا لمتوسط المجتمع لأن:

$$E(m) = \mu \quad , \qquad V(m) = \frac{\sigma^2}{n} \xrightarrow[n \to \infty]{} 0.$$