جامعة غليزان

كلية العلوم الإقتصادية، التجارية وعلوم التسيير

قسم علوم التسيير

المستوى: السنة الثانية شعبة علوم التسيير

مقياس: الاحصاء III (نظرية المعاينة)

الأستاذ مسؤول المقياس: د.حنصال أبوبكر

# محاضرات مقياس الاحصاء III (نظرية المعاينة)

السـنة الجامعية: 2021-2022

# المحاضرة الاولى:

## مفاهيم أساسية حول المتغيرة العشوائية والتوزيعات الاحتمالية

## 1. مفاهيم أساسية حول المتغيرة العشوائية والتوزيعات الاحتمالية

هناك عديد التوزيعات الاحتمالية المستمرة وغير المستمرة التي تخص المتغيرة العشوائية ولكن في هذا المحور سنتناول الا التوزيعات الأكثر شيوعا والتي لها أهمية كبيرة في التطبيقات الإحصائية التي سنتطرق لها في باقي المحاور لاسيما نظرية التقدير واختبار الفرضيات، حيث سنعطي الصيغة الرياضية الخاصة بكل توزيع وتدعيمنا بعديد الأمثلة التوضيحية عنه والتمثيالت البيانية الخاصة به.

### 1.1. مفهوم المتغيرة العشوائية المستمرة وتوزيعها الاحتمالي

**أ. تعريف المتغيرة العشوائية المستمرة**

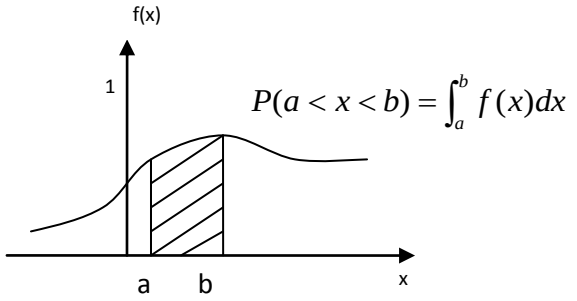
هي متغيرة $x_i$ تأخذ عددا لا متناهيا من القيم في مجال محدود أو تأخذ أي قيمة داخل هذا المجال، ولهذا فان وحدات قياس المتغيرة المستمرة تكون مستمرة مثل الزمن، الوزن، المسافة، الحجم، ...الخ (موراي، جون، و ألو،2004 ، ص.44).

**أ.1. التوزيع الاحتمالي المستمر**

هو مجموعة القيم التي يمكن أن تأخذها المتغيرة المستمرة $x_i$ والاحتمالات الملحقة بها، ونسمي توزيعا كهذا بدالة الكثافة الاحتمالية أو دالة الاحتمال، وتكون ممثلة بمنحنى متصل.

لاحظ أنه بما أن $x_i$ تأخذ عددا لا متناهيا من القيم داخل أي مجال، فإن احتمال قيمة بعينها هو احتمال يؤول إلى الصفر $P(X_i=x_i) \rightarrow 0$ لذلك فإن دالة الكثافة تستعمل لحساب احتمال مجال. ويكون ذلك بحساب المساحة تحت منحنى $f(x)$ بين حدود المجال (موراي، جون، و ألو، 2004 ، ص.44).

**الشكل.01: حساب الاحتمال من خلال شكل دالة الكثافة الاحتمالية**




**المصدر:** بوعبد الله صالح؛ **محاضرات الإحصاء الرياضي**، كلية العلوم الاقتصادية، جامعة المسيلة، الجزائر، السنة الجامعية 2005، 2006، ص.28

**أ.2. خصائص دالة الكثافة الاحتمالية للمتغيرة العشوائية المستمرة**

باستبدال إشارة التكامل بإشارة المجموع نجد أن شروط دالة الكثافة الاحتمالية للمتغيرة العشوائية المستمرة $\mathbf{x_i}$ تكتب كما يلي (موراي، جون، و ألو، 2004 ، ص.46):

$$1)\ f(x) \geq 0$$

$$2)\int_{-\infty}^{+\infty} f(x)dx = 1$$

من البديهي إذا أن منحنى دالة الكثافة لا يمكن أن ينزل أسفل محور المتغيرة العشوائية المستمرة ، كما أن المساحة الإجمالية بين المنحنى والمحور الأفقي تساوي الواحد، هذه الخصائص تفيدنا في حساب احتمالات بعض المجالات من خلال احتمالات مجالات أخرى.

**مثال.01:**

تأخذ دالة الكثافة الاحتمالية للمتغير العشوائي $\mathbf{x_i}$ الشكل التالي (موراي، جون، و ألو، 2004 ، ص.45):

$$f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & \text{sinon} \end{cases}$$

**المطلوب:**

- أوجد قيمة الثابت C التي تحقق الشرطين الأول والثاني لدالة الكثافة الاحتمالية
- أحسب احتمال أن تكون X تنتمي للمجال من 1 إلى 2.
- أحسب احتمال أن تكون X لا تنتمي للمجال من 1 إلى 2.

**الحل:**

لكي نقول عن دالة ما أنها دالة كثافة احتمالية يجب أن تكون موجبة دوما و أن يكون مجموع الاحتمالات مساويا للواحد.

- ايجاد قيمة الثابت C

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \Rightarrow \int_{-\infty}^{0} 0dx + \int_{0}^{3} Cx^2dx + \int_{3}^{+\infty} 0dx = 1 \Rightarrow C\left[\frac{x^3}{3}\right]_0^3 = 9C = 1 \Rightarrow C = 1/9$$

لكي تكون x دالة كثافة احتماية وتحقق الشرطين اعلاه يجب أن يكون C = 1/9 .

- حساب احتمال أن تكون X تنتمي للمجال من 1 إلى 2.

$$P(1 < x \le 2) = \int_1^2 f(x)dx = \int_1^2 (1/9)x^2dx = \frac{1}{9}\left[\frac{x^3}{3}\right]_1^2 = \frac{1}{9}\left[\frac{8-1}{3}\right] = \frac{7}{27}$$

- حساب احتمال أن تكون X لا تنتمي للمجال من 1 إلى 2.

$$P(1 > x > 2) = 1 - P(1 < x < 2) = 1 - \frac{7}{27} = \frac{20}{27}$$

### أ.3. دالة التوزيع F(x) للمتغيرة العشوائية المستمرة

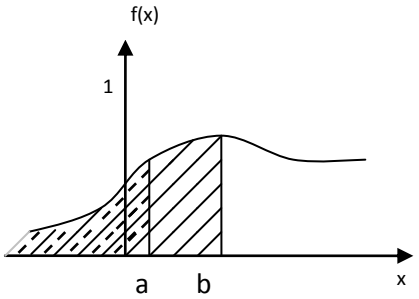
تعرف دالة التوزيع للمتغيرة المستمرة كما يلي (موراي، جون، و ألو، 2004 ، ص.46):

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)du$$

ان لدالة التوزيع أهمية أكبر بالنسبة للمتغيرة المستمرة والسبب في ذلك أننا نهتم في حالة المتغيرة المستمرة باحتمال مجال وليس باحتمال نقطة، ولحساب احتمال مجال من الأيسر التعويض في دالة التوزيع بدلا من حساب التكامل في كل مرة. يتضح ذلك من القاعدة التالية:

بفرض a, b نقطتان من مجال تعريف $\mathbf{x_i}$، بحيث b > a .

**الشكل.02: شكل دالة التوزيع التراكمي**




**المصدر:** بوعبد الله صالح؛ **محاضرات الإحصاء الرياضي**، كلية العلوم الاقتصادية، جامعة المسيلة، الجزائر، السنة الجامعية 2005، 2006، ص.28

لحساب احتمال أن تكون $\mathbf{x_i}$ تنتمي إلى المجال ]a, b[ :

نكتب ما يلي: $P(a < x \le b) = P(X \le b) - P(X \le a) = F(b) - F(a)$

**مثال.02**:

أوجد دالة التوزيع F(x) للمتغيرة المذكورة في المثال السابق، ثم استخدم دالة التوزيع لحساب الاحتمال: P(1< x <2).

- ايجاد شكل دالة التوزيع F(x)

$$*\ x < 0: \qquad F(x) = \int_{-\infty}^{x} f(u)du = \int_{-\infty}^{0} 0du = 0$$

$$*\ 0 \le x < 3: \qquad F(x) = \int_{-\infty}^{x} f(u)du = \int_{-\infty}^{0} 0du + \int_{0}^{x} \frac{1}{9}u^2 du = \frac{1}{9}\left[\frac{u^3}{3}\right]_0^x = \frac{x^3}{27}$$

$$*\ x \ge 3: \qquad F(x) = \int_{-\infty}^{x} f(u)du = \int_{-\infty}^{0} 0du + \int_{0}^{3} \frac{1}{9}u^2 du + \int_{3}^{x} 0du = 0 + \frac{1}{9}\left[\frac{u^3}{3}\right]_0^3 + 0 = \frac{27}{27} = 1$$

ومنه فان دالة التوزيع F(x) تكتب على الشكل التالي:

$$F(x) = \begin{cases} 0 & x < 0 \\ x^3 / 27 & 0 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$

- حساب الاحتمال: P(1< x <2) باستخدم دالة التوزيع

$$P(1 < x \le 2) = F(2) - F(1) = \frac{2^3}{27} - \frac{1^3}{27} = \frac{7}{27}$$

**أ.4. قاعدة لايبنيز**

تفيد هذه القاعدة الرياضية العامة في استنتاج أن مشتقة دالة التوزيع هي عبارة عن دالة كثافة احتمالية، وهي كما يلي (صالح، 2006 ، ص.30):

$$\frac{d\int_{-\infty}^{x} f(u)du}{dx} = f(x) \Rightarrow \frac{dF(x)}{dx} = f(x)$$

**مثال.03**:

أوجد دالة الكثافة الاحتمالية للمتغيرة X إذا كانت دالة التوزيع كما يلي:

$$F(x) = \begin{cases} 1-e^{-2x} & x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

لايجاد دالة الكثافة الاحتمالية للمتغيرة X يجب البحث عن مشتقة دالة التوزيع الخاصة بنفس المتغيرة، وذلك كما يلي:

$$* x < 0 : f(x) = F'(x) = (0)' = 0$$

$$* x \geq 0 : F(x) = 1-e^{-2x} \Rightarrow f(x) = \left(1-e^{-2x}\right)' = 2e^{-2x}$$

ومنه فان دالة الكثافة الاحتمالية للمتغيرة X تعطى بالشكل التالي:

$$f(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & \text{sinon} \end{cases}$$

### 2.1. التوقع الرياضي والتباين:

في العديد من الحالات لا يكفي حساب احتمال تحقق حدث أو أحداث معينة بل نحتاج للخروج بتوقع معين يلخص الوضعية المطروحة أمامنا. من جهة أخرى قد يصعب المفاضلة بين خيارات متاحة مقيمة بمبالغ معينة بسبب ارتباط كل مبلغ بمخاطرة مختلفة؛ من المعروف أن الاستثمارات الأكثر مردودية هي تلك التي تتضمن أكبر مخاطرة، فكيف يمكن أخذ في الحسبان المخاطرة والمبلغ المتوقع وبطريقة دقيقة وموضوعية؛ إن طريقة التوقع وبقية المفاهيم الأخرى الواردة أعلاه يمكن أن تساعدنا في ذلك.

#### 1.2.1. التوقع الرياضي (l'Espérance mathématique)

**أ. تعريف:**

يعرف التوقع الرياضي لمتغيرة عشوائية متقطعة كما يلي (موراي، جون، و ألو، 2004 ، ص.36):

$$E(X) = \sum_{i=1}^{n} x_i P(X = x_i) = \sum_{i=1}^{n} x_i f(x_i)$$

ويعرف التوقع الرياضي لمتغيرة عشوائية مستمرة كما يلي:

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

كما نرمز للتوقع أحيانا بالرمز $\mu$ أو $\mu_X$.

**مثال.04:**

أحسب العدد المتوقع من المرات التي نحصل فيها على وجه قطعة نقدية عند قيامنا بتجربة رميها لـ 4 مرات.

**الحل:**

- **جدول التوزيع الاحتمالي:**

| عدد مرات وجه **X** | 0 | 1 | 2 | 3 | 4 | المجموع |
|---|---|---|---|---|---|---|
| P(X) | $(1/2)^4$ | $4\ (1/2)^4$ | $6\ (1/2)^4$ | $4\ (1/2)^4$ | $(1/2)^4$ | 16/16 = 1 |
| XP(X) | 0 | $4\ (1/2)^4$ | $12\ (1/2)^4$ | $12\ (1/2)^4$ | $4\ (1/2)^4$ | 32/16 = 2 |

- **حساب التوقع الرياضي:**

بالتعويض في المعادلة أعلاه نجد:

$$E(X) = \sum_{i=1}^{n} x_i P(X = x_i) = 32/16 = 2$$

اذا العدد المتوقع هو مرتين من بين 4 رميات.

**مثال.05**:

نلقي زهرة نرد مرة واحدة، يربح اللاعب 20 دج إذا حصل على الرقم 2، ويربح 40 دج إذا حصل على الرقم 4، و60 دج إذا حصل على الرقم 6، ويخسر 10 دج إذا حصل على الرقم 1، 30 دج إذا حصل على الرقم 3، و50 دج إذا حصل على الرقم 5.

**المطلوب:**

تحقق مما إذا كانت اللعبة متوازنة (هل توقع الربح يساوي توقع الخسارة).

**الحل:**

- **تحديد جدول التوزيع الاحتمالي وحساب التوقع**

| **نتيجة الرمي** | 1 | 2 | 3 | 4 | 5 | 6 | المجموع |
|---|---|---|---|---|---|---|---|
| **نتيجة المراهنة X** | -10 | 20 | -30 | 40 | -50 | 60 | - |
| **P(X=x)** | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 6/6 |
| **X*P(X=x)** | -10/6 | 20/6 | -30/6 | 40/6 | -50/6 | 60/6 | (120-90)/6>0 |

$$E(x) = 30/6 = 5 > 0$$

اذا نستنتج أن اللعبة غير متوازنة لأن توقع الربح أكبر من توقع الخسارة

**مثال.06:**

أوجد التوقع الرياضي للمتغيرة ذات دالة الكثافة الاحتمالية التالية:

$$f(x) = \begin{cases} \frac{1}{2}x & 0 < x < 2 \\ 0 & \text{sinon} \end{cases}$$

$$E(x) = \int_{-\infty}^{\infty} x f(x)dx = \int_{-\infty}^{0} x.0.dx + \int_{0}^{2} x(\frac{1}{2}x)dx + \int_{2}^{\infty} x.0.dx$$

**الحل:**

التوقع الرياضي يحسب وفق العلاقة التالية:

$$E(x) = 0 + \frac{1}{2}\left[\frac{x^3}{3}\right]_0^2 + 0 = \frac{4}{3}$$

**تمرين:**

أرسلت مؤسسة عروضا إلى 4 عملاء. احتمال تلقي طلبية من العميل الأول هي 0.2، من العميل الثاني 0.3، من العميل الثالث 0.35 و0.4 من العميل الرابع.

**المطلوب:**

في انتظار ردود العملاء ما هو العدد المتوقع من الطلبيات؟

**ب. توقع دالة**

يستخدم توقع دالة عند حساب عدد من المقاييس مثل التباين، العزوم المركزية والعزوم المرتبطة بالأصل.

لتكن X متغيرة عشوائية لها دالة كثافة f(x)، و$y = g(x)$ متغيرة عشوائية تابعة لها، تكتب دالة التوقع كما يلي (موراي، جون، و ألو، 2004 ص.48):

- ✓ في حالة X متغيرة عشوائية منفصلة:

$$E(Y) = E(g(x)) = \sum_x g(x) f(x)$$

- ✓ في حالة X متغيرة عشوائية متصلة:

$$E(Y) = E(g(x)) = \int_{-\infty}^{+\infty} g(x) f(x) dx$$

**مثال.07**. نلقي قطعة نقدية مرتين، نسمي X عدد مرات الحصول على صورة،

**المطلوب :**

أحسب E(X) وE(Y) اذا علمت أن Y تحسب وفق العلاقة التالية:

$Y=X^2$

| X | 0 | 1 | 2 | المجموع |
|---|---|---|---|---|
| P(X = x) | 1/4 | 1/2 | 1/4 | 1 |
| X*P(X) | 0 | 1/2 | 1/2 | E(X) =1 |
| $X^2$ | 0 | 1 | 4 | |
| $X^2$*P(X) | 0 | 1/2 | 1 | $E(X^2) = 3/2$ |

**مثال.08**:

لتكن $x_i$ متغيرة عشوائية ذات دالة الكثافة التالية:

$$f(x)=\begin{cases}\frac{1}{2}x & 0<x<2\\ 0 & \text{sinon}\end{cases}$$

**المطلوب :**

أحسب E(Y) اذا علمت أن Y تحسب وفق العلاقة التالية:

$$Y = g(x) = 3x^2 - 2x$$

**الحل:**

$$E(Y)=E(g(x))=\int_{-\infty}^{+\infty}g(x)f(x)dx=\int_{-\infty}^{+\infty}(3x^2-2x)(0)dx+\int_{-\infty}^{+\infty}(3x^2-2x)(x/2)dx+\int_{-\infty}^{+\infty}(3x^2-2x)(0)dx$$

$$E(Y)=0+\int_0^2(3x^2-2x)(x/2)dx+0=\frac{1}{2}\left[\frac{3x^4}{4}-\frac{2x^3}{3}\right]_0^2=\frac{1}{2}\left[12-\frac{16}{3}\right]=\frac{20}{6}=\frac{10}{3}$$

**ج. خصائص التوقع الرياضي**

سنذكر بايجاز أهم خصائص التوقع الرياضي كما يلي (صالح، 2006 ص.33):

i. E (C) = C
ii. E(CX) = CE(X)
iii. E(X+Y) = E(X) + E(Y)
iv. E(XY) = E(X)E(Y) إذا كانت المتغيرتان مستقلتان.

**مثال.09:**

تتوقع مؤسسة أن تتلقى كل شهر 3 طلبيات من العميل A و4 من B .

- أحسب العدد المتوقع من الطلبيات المتلقاة من A في السنة.
- أحسب العدد الإجمالي المتوقع من الطلبيات المتلقاة في الشهر.

يعين كل عميل من طرفه مندوبا عن كل طلبية لمتابعة إتمامها، كم تتوقع أن يلزم من مقابلة لتعريف مندوبي العميل A بمندوبي B.

**الحل:**

E(12A) = 12E(A) = 12(3) = 36
E(A*B) = E(A)*E(B) = 4*3 = 12.
E(A + B) = E(A) + E(B) = 3 + 4 = 7

**2.2.1. التباين**

**أ. تعريف:**

يعرف التباين لمتغيرة عشوائية وفق العلاقة التالية (موراي، جون، و ألو، 2004 ص.47):

$$\sigma_X^2 = Var(X) = E\left[(X-\mu)^2\right]$$

والانحراف المعياري هو جذر التباين.

$$\sigma = \sqrt{V(X)} = \sqrt{\sigma^2}$$

في حالة المتغيرة العشوائية المتقطعة :

$$V(X) = \sum_x (x-\mu)^2 f(x)$$

في حالة المتغيرة العشوائية مستمرة :

$$V(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx$$

**مثال.10 :**

نلقي قطعة نقدية مرتين. نسمي Xi عدد مرات الحصول على صورة، أحسب V(X) .

| X | 0 | 1 | 2 | المجموع |
|---|---|---|---|---|
| P(X = x) | 1/4 | 1/2 | 1/4 | 1 |
| X*P(X) | 0 | 1/2 | 1/2 | E(X) = μ =1 |
| $(X-\mu)^2$ | 1 | 0 | 1 | |
| $(X-\mu)^2$ * P(X) | 1/4 | 0 | 1/4 | V(X) = 1/2 |

**مثال.11:**

لتكن X م ع ذات دالة الكثافة التالية :

$$f(x) = \begin{cases} \frac{1}{2}x & 0 < x < 2 \\ 0 & \text{sinon} \end{cases}$$

أحسب تباين X. اذا علمت أن ($\mu$=4/3 )

**الحل:**

نعلم أن :

$$\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx. \qquad \mu = 4/3$$

بالتعويض نجد:

$$\sigma^2 = \int_{-\infty}^{0} \left(x-\frac{4}{3}\right)^2 * 0dx + \int_0^2 \left(x-\frac{4}{3}\right)^2 \frac{x}{2}dx + \int_2^{+\infty} \left(x-\frac{4}{3}\right)^2 0dx = 0 + \frac{1}{2}x\left(x^2 - \frac{8x}{3} + \frac{16}{9}\right)_0^2 + 0$$

$$\sigma^2 = \frac{1}{2}\left(x^3 - \frac{8x^2}{3} + \frac{16x}{9}\right)_0^2 = \frac{1}{2}\left(8 - \frac{32}{3} + \frac{32}{9}\right) = \frac{4}{9}$$

**ب. خصائص التباين** (صالح، 2006 ص.34):

i. $V(X) = E(X^2) - E(X)^2$
ii. $V(CX) = C^2V(X)$ , $V(C) = 0$
iii. $V(X + Y) = V(X) + V(Y)$
iv. في حالة استقلال المتغيرتان عن بعضهما $V(X - Y) = V(X) + V(Y)$

**مثال.12:**

نلقي قطعة نقدية مرتين. نسمي X عدد مرات الحصول على صورة، أحسب V(X) باستخدام الصيغة

$$V(X) = E(X^2) - E(X)^2 .$$

لتكن المتغيرة $Y = 2X$.

أحسب V(Y)

نلقي حجر نرد ونسمي Z النتيجة المحصل عيها.

أحسب تباين المتغيرة W حيث: $W = Z - Y$

**الحل:**

| X | 0 | 1 | 2 | المجموع |
|---|---|---|---|---|
| P(X = x) | 1/4 | 1/2 | 1/4 | 1 |
| X P(X) | 0 | 1/2 | 1/2 | E(X) = μ =1 |
| $(X)^2$ | 0 | 1 | 4 | |
| $(X)^2 * P(X)$ | 0 | 1/2 | 1 | $E(X)^2 = 3/2$ |

$V(X) = E(X^2) - E(X)^2 = (3/2) - 1 = 1/2$

$V(Y) = V(2X) = 2^2V(X) = 4\ (1/2) = 2.$

$V(W) = V(Z-2X) = V(Z) +V(2X) = V(Z) + 2^2V(X).$

$V(Z) = E(Z^2) – E(Z)^2 = (1/6)[1^2+2^2+3^2+4^2+5^2+6^2]-(1/6)[1+2+3+4+5+6] = 70/6$

$V(W) = (70/6) + 4\ (1/2) = 82/6 = 13.67$

### 3.1. المتغيرة المعيارية Variable centrée réduite

يمكن أن نلحق بأي متغيرة عشوائية X متغيرة معيارية (تسمى أيضا المتغيرة المركزية) ويرمز لها بالرمز X* ، تلحق المتغيرة المعيارية بالمتغيرة الحقيقية من أجل المقارنة لأن المتغيرة المعيارية ليس لها وحدة كالمتر أو الساعة ...الخ وإنما هي تعبر عن كل قيمة x لـ X من خلال المسافة بين x والتوقع μ ليست محسوبة بالوحدة الأصلية وإنما بالانحرافات المعيارية (صالح، 2006 ، ص.34).

$$X^* = \frac{X - \mu}{\sigma}$$

فمن خلال خصائص التوقع والتباين نستخرج التوقع والتباين للمتغيرة المعيارية.

$$E(X^*) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}E(X-\mu) = \frac{1}{\sigma}\left[E(X) - E(\mu)\right] = 0$$

$$V(X^*) = E\left[(X^* - E(X^*))^2\right] = E\left[(X^* - 0)^2\right] = E(X^{*2}) = E\left(\frac{(X-\mu)^2}{\sigma^2}\right) = \frac{1}{\sigma^2}E(X-\mu)^2 = \frac{\sigma^2}{\sigma^2} = 1$$

**مثال.13**:
أحسب *X من أجل متوسط 70، انحراف معياري 5، وXi يساوي: 55، 60، 50، 75، 80، 70.
**الجواب:**
القيم هي: -3، -2، -4، 1، 2، 0.

**مثال.14:**
يتدرب عاملان أحمد وعلي من أجل المشاركة في ماراتون عيد العمال، حيث يشترط يوم المسابقة أن لا يتجاوز وزن المترشح المجال $\mu \pm 1.5\sigma$ .
**المطلوب:**
إذا كان الوزن المتوسط بالكغ هو $\mu = 70$ والانحراف المعياري هو 5 كغ، هل سيقبل العاملان أحمد وعلي إذا كان وزنهما: 77 كغ، و80 كغ؟
**الحل:**
مجال القبول هو من 62.5 كغ إلى 77.5 كغ، لذلك فسيرفض علي ويقبل أحمد.

**كخلاصة** نقول أنه من أجل المقارنة بين المتغيرات تستخدم المتغيرة المعيارية التي تسمح بالتعبير عن قيمة x ليس من خلال وحداتها الأصلية (كغ، متر، زمن، ...) وإنما بعدد الانحرافات المعيارية التي تفصل بين القيمة x والتوقع الرياضي.
التوقع الرياضي والتباين للمتغيرة المعيارية $X^* = \frac{X-\mu}{\sigma}$ هما على التوالي 0 و 1.
لحساب التوقع الرياضي لدالة ما في X (مثلا التباين، أو $X^2$) نضرب قيم الدالة في الاحتمالات المقابلة لـ X:

$$E(Y) = E(g(x)) = \sum_x g(x) f(x)$$

$$E(Y) = E(g(x)) = \int_{-\infty}^{+\infty} g(x) f(x) dx$$

# المحاضرة الثانية:

## بعض التوزيعات الاحتمالية المستمرة الأكثر شيوعا

## 2.1. بعض التوزيعات الاحتمالية المستمرة الأكثر شيوعا

سنتناول في هذا العنصر من المحور الأول للمطبوعة عدد من التوزيعات ذات الاستخدام الواسع في الإحصاء الاستدلالي والتطبيقي وبخاصة في مجالي "التقدير" و"اختبار الفروض"، حيث يستخدم الإحصائي هذه التوزيعات في عمله من أجل الوصول إلى "قرار" بشأن "المجتمع" المدروس انطلاقا من بيانات يتحصل عليها من عينة، ونجد من بين أهم هذه التوزيعات ما يلي:

### 2.1. 1. التوزيع الطبيعي أو توزيع لابلاس قوس - D. Normale ou D. de Laplace Gausse

سمي هذا التوزيع نسبتا الى العالمان الرياضيان الفيزيائيان والفلكيان الفرنسي Pière Simon de Laplace (1749-1827) والألماني Carl Freidrich Gauss، (1777-1855)الذين كانا من أوائل من اكتشف هذا القانون. أما من أعطاه تسمية التوزيع الطبيعي فهو Pearson في 1893. (دروزبيك، 1996، ص.329)؛ حيث يعتبر التوزيع الطبيعي من أهم التوزيعات الاحتمالية شائعة الاستخدام لما له من خصائص تنطبق على نسبة كبيرة من الظواهر الطبيعية والاجتماعية والاقتصادية؛ فلو اخترنا بالصدفة مئة أو ألفا من المارين في شارع ما وقسنا أطوالهم لوجدنا نسبة كبيرة منها قريبة من متوسط ما، ونسبة قليلة من طوال القامة ونسبة مقاربة لها من قصار القامة. ومثل هذا بالنسبة للأوزان. ولو مثلنا هذه البيانات في معلم متعامد متجانس لكان المنحنى الذي يمثل النسبة، أو ما يمكن أن نسميه الاحتمال، ذا شكل جرسي متماثل حول المتوسط وهي صفات التوزيع الطبيعي (محمد حسين، 2008، ص.227) (الشكل.03) :

**أ. صيغة القانون**

تكتب دالة الكثافة لمنحنى للتوزيع الطبيعي كما يلي:

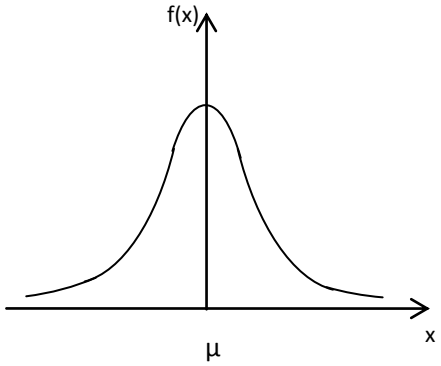
$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-(1/2)\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty<x<\infty$$

حيث $\mu$ و$\sigma$ هما على التوالي التوقع والانحراف المعياري؛

ونكتب $X \sim N(\mu, \sigma)$

حيث يمثل الشكل العام لمنحنى للتوزيع الطبيعي كما يلي:

**الشكل .03: الشكل العام للتوزيع الطبيعي**




**المصدر:** بوعبد الله صالح؛ **محاضرات الإحصاء الرياضي**، كلية العلوم الاقتصادية، جامعة المسيلة، الجزائر، السنة الجامعية 2005، 2006، ص.52

وتكتب دالة التوزيع (الدالة التجميعية) للتوزيع الطبيعي تكتب كما يلي (صالح، 2006، ص.52):

$$F(x) = P(X \le x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-1/2\left(\frac{v-\mu}{\sigma}\right)^2} dv$$

**ب. المتغيرة المركزية أو المعيارية** (صالح،2006 ، ص.ص، 52.53)**:**

تستخدم المتغيرة المعيارية $Z = (X-\mu)/\sigma$ لتكوين الجداول الإحصائية للاحتمالات:

$$F(z) = P(Z \le z) \text{ أو } P(0 \le Z \le z)$$

حيث تسمح بكتابة الدالة f وF بدلالة مجهول واحد Z بدلا من 3 مجاهيل x وμ و σ وذلك كما يلي:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \qquad -\infty < z < \infty$$

والدالة التجميعية بـ:

$$F(z) = P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-u^2}{2}} du$$

بالنظر إلى العلاقة الخطية بين المتغيرتين X وZ، فإن Z تتبع نفس توزيع X أي التوزيع الطبيعي. ونعلم أن:

$$\mathbf{V(Z) = 1} \qquad \mathbf{E(Z) = 0}$$

**ج. خصائص التوزيع الطبيعي** (صالح، 2006، ص.53) **:**

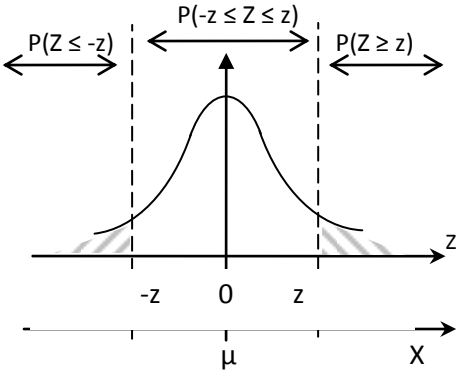
من خصائص التوزيع الطبيعي أنه يعتبر معتدلا لا مدببا ولا مفلطحا، حيث يعتبر معامل التفلطح $\alpha_4 = 3$ للتوزيع الطبيعي معيارا لاعتدال المنحنيات.

كذلك من خصائصه أنه متماثل حول القيمة المتوقعة $\alpha_3 = \frac{\mu_3}{\sigma^3} = 0$

**تماثل** منحنى X حول المتوسط (أنظر الشكل 3) يعني تماثل لمنحنى Z حول 0، مما يعني أنه من أجل أي قيمة للمتغيرة المعيارية $z > 0$ (محمد حسين،2008، ص.229):

$$P(0 \leq Z \leq z) = P(-z \leq Z \leq 0) = P(-z \leq Z \leq z) / 2$$
$$P(Z \leq -z) = 1 - P(Z \leq z) = P(Z \geq z)$$

**الشكل.04: استخدام تماثل التوزيع الطبيعي في حساب الاحتمالات**



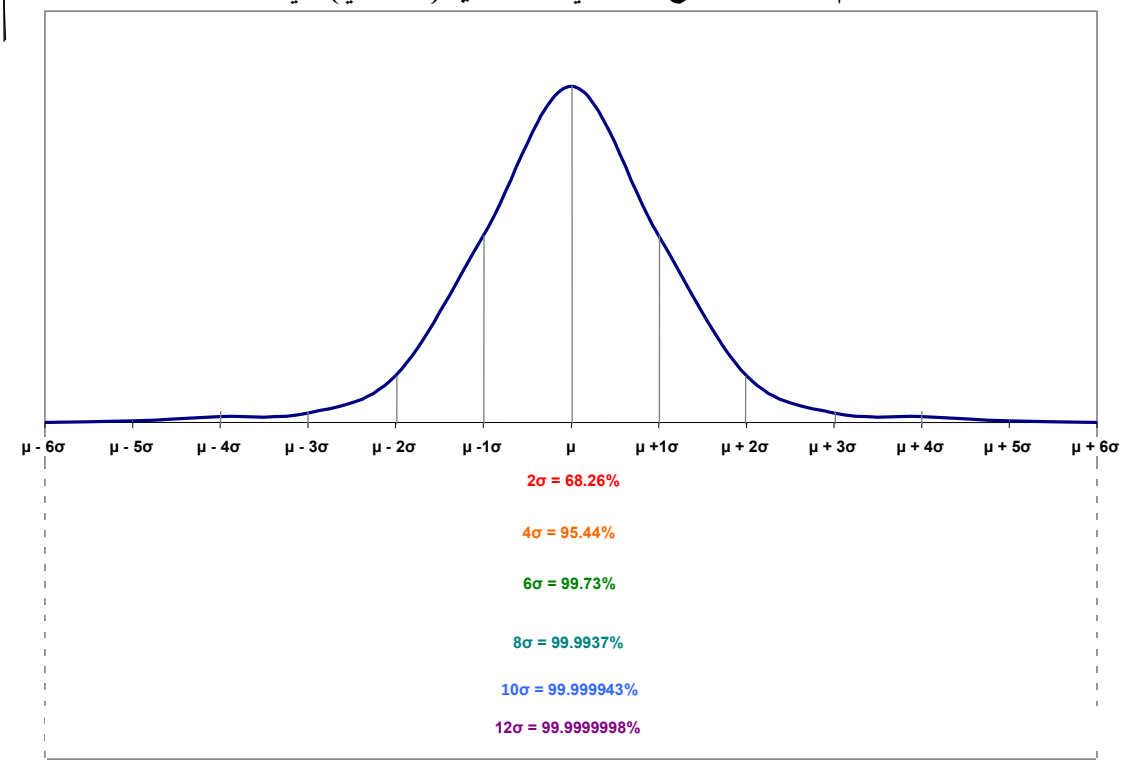

**المصدر:** بوعبد الله صالح؛ **محاضرات الإحصاء الرياضي**، كلية العلوم الاقتصادية، جامعة المسيلة، الجزائر، السنة الجامعية 2005، 2006، ص.53

لقد تم باستخدام المتغيرة المعيارية Z حساب الاحتمالات (المساحات) تحت المنحنى ومنها خاصة:

$$P(-\sigma \leq X \leq \sigma) = P(-\sigma \leq Z \leq \sigma) = 0.6837$$
$$P(-2\sigma \leq X \leq 2\sigma) = P(-2 \leq Z \leq 2) = 0.9544$$
$$P(-3\sigma \leq X \leq 3\sigma) = P(-3 \leq Z \leq 3) = 0.9973$$

هذه القيم وغيرها متوفرة في الجداول الإحصائية التي نجدها في الكثير من المراجع.

**الشكل.05: استخدام تماثل التوزيع الطبيعي المعياري (القياسي) في حساب الاحتمالات**



**المصدر:** موراي شبيجل؛ جون شيلر؛ ألو سرينيقاسان، ترجمة محمود علي أبو النصر ومصطفى جلال مصطفى، **ملخصات ايزي شوم، الاحتمالات والاحصاء**، الدار الدولية للاستثمارات الثقافية، مصر، الطبعة الأولى، 2004، ص.55

**مثال.15**: باستعمال الجداول الاحصائية

1) أحسب : $P(0 \leq Z \leq z)$ من أجل $z = 1, 2, 3$

2) أحسب $P(-z \leq Z \leq z)$ من أجل نفس القيم لـ z.

**الحل:**

1) 0.3413 ، 0.47725 ، 0.49865

2) 0.6827، 0.9545، 0.9973

## 2.1.2. توزيع ك2 Distribution en Khi-carré (ou Khi-deux)

### أ. صيغة القانون

توزيع ك2 يعتبر من بين أكثر التوزيعات استخداما في مجال اختبارات الفروض بأنواعها، ويمكن تعريفه كما يلي (دومينيك، 2011، ص.ص، 136.137):

اذا كانت $X_1, X_2, \ldots X_v$ ، متغيرات عشوائية مستقلة كل منها تتبع التوزيع الطبيعي المعياري $(\mu = 0, \sigma = 1)$.

ولتكن المتغيرة العشوائية X التالية والتي تكتب كما يلي:

$$X = X_1^2 + X_2^2 + \ldots + X_v^2$$

والتي لها دالة الكثافة الاحتمالية التالية:

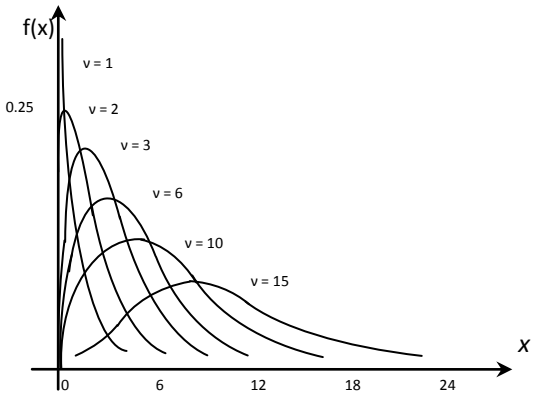
$$f(x) = \begin{cases} \dfrac{x^{(v/2)-1} e^{-x/2}}{2^{v/2}\Gamma(v/2)} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

حيث $\Gamma(\alpha)$ هي الدالة قاما:

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

والشكل العام لمنحنى توزيع ك2 يمثل كما يلي:

**الشكل.06: تدرج منحنى ك2 حسب درجة الحرية**




**المصدر:** بوعبد الله صالح؛ **محاضرات الإحصاء الرياضي**، كلية العلوم الاقتصادية، جامعة المسيلة، الجزائر، السنة الجامعية 2005، 2006، ص.69

ونقول أن X تتبع التوزيع ك2 ب ν درجة حرية ونكتب: $X \sim \chi^2_{\nu}$ .

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

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

**ب. خصائص توزيع ك 2** (صالح، 2006، ص. 70):

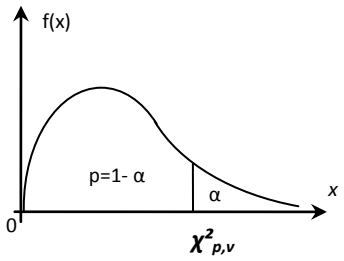
$$E(X) = \nu, \quad V(X) = 2\nu, \quad M(t) = (1-2t)^{-\nu/2}$$

دالة التوزيع ك2 هي حالة خاصة من توزيع قاما بوضع $\alpha = \nu/2, \beta = 2$

ويأخذ منحنى f(x) شكله حسب قيمة الثابت $\nu$ ونلاحظ من الرسم أن المنحنى يبتعد شيئا فشيئا عن المحور العمودي ويأخذ شكلا جرسيا كلما زادت قيمة $\nu$. ونبرهن أنه عند $\nu$ كبير ($\nu \ge 30$) فإن $\sqrt{2\chi^2} - \sqrt{2\nu - 1}$ تتبع التوزيع الطبيعي المعياري.

- **في الجداول الاحصائية**، تعين نقطة (قيمة المتغيرة) ك2 على المحور الأفقي (أنظر الرسم المقابل) من خلال $\nu$ بالإضافة إلى المساحة p على يسار ك2 تحت المنحنى ($p = P(X \le \chi^2_{\nu;p})$). وأحيانا تحدد النقطة ك2 بدلالة المساحة على يمينها ($\alpha = 1 - p$) لذلك نجد في كتب الاحصاء كل من الكتابتين: $\chi^2_{p,\nu}$ و $\chi^2_{\alpha,\nu}$

**الشكل.07: تعيين نقطة ك2 على المحور من خلال قيمة p**




**المصدر:** بوعبد الله صالح؛ **محاضرات الإحصاء الرياضي**، كلية العلوم الاقتصادية، جامعة المسيلة، الجزائر، السنة الجامعية 2005، 2006، ص.70

نعبر عن قيمة المتغير العشوائي الذي يتبع توزيع مربع كاي بالرمز $\chi^2_{(\nu;p)}$. وهي القيمة التي يقع على يمينها مساحة $\alpha$ على منحنى كاي تربيع بدرجة حرية $\vartheta$ وتحسب قيم هذا المتغير العشوائي والاحتمالات الموافقة له باستخدام الجدول الاجتمالي الخاص بهذا التوزيع والمرفق في قائمة الملاحق، والمثال التالي يوضح كيفية حساب قيم الاحتمالات بالاعتماد على هذا الجدول.

**مثال.16 :**

اوجد القيم التالية : $\chi^2_{(0.1;17)}$ $\chi^2_{(0.995;8)}$ $\chi^2_{(0.95;7)}$ ;

**الحل:**

لإيجاد $\chi^2_{(0.95;7)}$ نختار من العمود الايسر، عمود درجة الحرية القيمة 7 فنجد على يمينه صف من الأعداد ومن الصف العلوي (صف المساحات أي الاحتمالات)، عند القيمة 95.0 فنجد أسفلها عمود من الاعداد فتكون القيمة الموجودة في تقاطع الصف والعمود هي القيمة المطلوبة وبالتالي يكون:

$$\chi^2_{(0.95;7)} = 2.167$$

$$\chi^2_{(0.1;17)} = 24.767$$

$$\chi^2_{(0.995;8)} = 1.344$$

### 2.1. 3. توزيع ستيودنت Distribution de Student

**أ. صيغة القانون**

لتكن المتغيرتان العشوائيتان المستقلتان Y وZ حيث $Y \sim N(0, 1)$ و $Z \sim \chi^2_\nu$ ؛ المتغيرة

$$T = \frac{Y}{\sqrt{Z/\nu}}$$

تتبع توزيع ستيودنت ولها دالة الكثافة التالية (دروزبيك، 1996، ص.262):

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

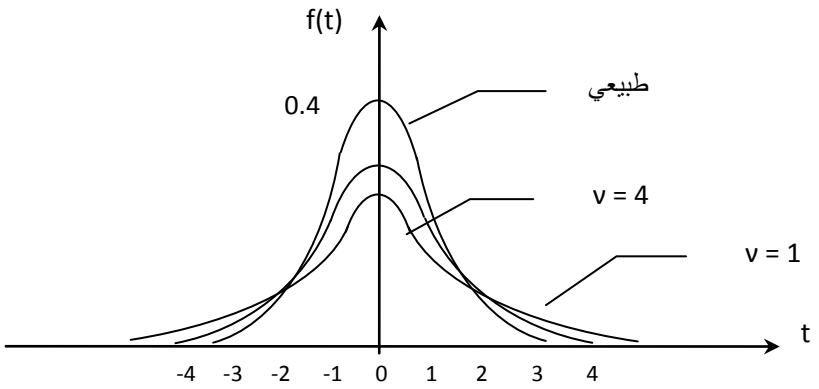
$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt \qquad \alpha > 0$$

ونقول أن المتغيرة X تتبع توزيع ستيودنت ب $\nu$ درجة حرية ونكتب: $T \sim t_\nu$

**ب. خصائص توزيع ستيودنت**

$$E(T) = 0, \qquad V(T) = \nu/(\nu-2) \quad si \quad (\nu > 2)$$

**الشكل.08: تدرج منحنى ستيودنت حسب درجة الحرية**




**المصدر:** موراي شبيجل؛ جون شيلر؛ ألو سرينيقاسان، ترجمة محمود علي أبو النصر ومصطفى جلال مصطفى، **ملخصات ايزي شوم، الاحتمالات والاحصاء**، الدار الدولية للاستثمارات الثقافية، مصر، الطبعة الأولى، 2004؛ ص.55

- نلاحظ أن منحنى t متماثل حول المتوسط 0 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 $\nu$ . وعموما، يعتبر الإحصائيون أن المنحنيان يتطابقان تقريبا عند $\nu \geq 30$.
- في الجداول الاحصائية، تعين نقطة (قيمة المتغيرة) t من خلال $\nu$والمساحة p على يسار t تحت المنحنى

( $p = P(T \leq t_{\nu;p})$). وأحيانا تحدد النقطة t بدلالة المساحة على يمينها ($\alpha = 1 - p$) ونكتب : $t_{p,\nu}$ أو $t_{\alpha,\nu}$

نعبر عن قيمة المتغير العشوائي الذي يتبع توزيع Student بالرمزt وهي القيمة التي تقع على يمينيا مساحة $\alpha$ وعلى منحى توزيع t بدرجة حرية $\vartheta$، وتحسب قيم هذا المتغير العشوائي والاحتمالات الموافقة له باستخدام الجدول الاحتمالي الخاص بهذا التوزيع والمرفق في قائمة الملاحق، والمثال التالي يوضح كيفية حساب قيم الاحتمالات بالاعتماد على هذا الجدول.

**مثال.17 :**

اوجد القيم التالية : $t_{(0.995;9)}$ $t_{(0.99;11)}$ $t_{(0.05;3)}$

**الحل:**

لايجاد $t_{(0.05;3)}$ نختار من العمود الايسر (عمود درجة الحرية) القيمة 3 فنجد على يمينه صف من الاعداد ومن الصف العلوي (صف المساحات أي الاحتمالات) عند القيمة 0.05 فنجد

أسفلها عمود من الاعداد فتكون القيمة الموجودة في تقاطع الصف والعمود هي القيمة المطلوبة وبالتالي يكون :

$$\mathbf{t}_{(0.05;3)} = 2.353$$

الملاحظ أنه لايجاد $\mathbf{t}_{(0.99;11)}$ فان جدول التوزيع لا يحتوي على القيمة $\alpha=0.99$ لذلك نستغل تماثل منحنى التوزيع حول الصفر فيكون:

$$\mathbf{t}_{(0.99;11)} = -\mathbf{t}_{(0.01;11)} = -2.718$$

وبالمثل يكون:

$$\mathbf{t}_{(0.995;9)} = -\mathbf{t}_{(0.005;9)} = -3.250$$

## 2.1. 4. توزيع فيشر Distribution F de Fisher-Snédecor (F)

### أ. صيغة القانون (Michel، 2010 .p.76):

ليكن لدينا المتغيرتان العشوائيتان المستقلتان $X_1 \sim \chi^2_{\nu_1}$ و $X_2 \sim \chi^2_{\nu_2}$ . فان النسبة :

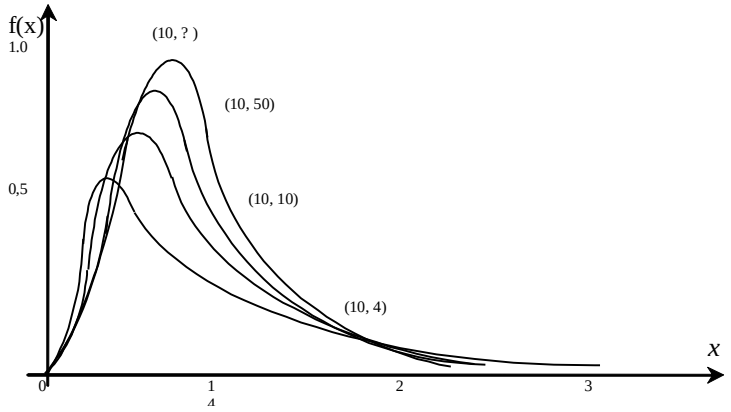
$$F = \frac{X_1 / \nu_1}{X_2 / \nu_2}$$

بين التوزيعين تعرف بتوزيع F بالمعلمتين $\vartheta_1$ و $\vartheta_2$، وهما درجة حرية البسط والمقام على الترتيب، واقتران دالة الكثافة الاحتمالية للتوزيع تعطى كما يلي:

$$f(x) = \begin{cases} \dfrac{\Gamma\left(\dfrac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\dfrac{\nu_1}{2}\right)\Gamma\left(\dfrac{\nu_2}{2}\right)} \; \nu_1^{\frac{\nu_1}{2}} \; \nu_2^{\frac{\nu_2}{2}} \; x^{(\frac{\nu_1}{2})-1} \; \left(\nu_2+\nu_1 x\right)^{-\frac{\nu_1+\nu_2}{2}} & x > 0 \\ 0 & x \le 0 \end{cases}$$

ونقول أن المتغيرة X تتبع توزيع فيشر بـ $\nu_1$ و$\nu_2$ درجة حرية ونكتب: $X \sim F_{\nu 1,\, \nu 2}$

**الشكل.09: تدرج منحنى فيشر حسب درجة الحرية**




Source : Michel Lejeune **, Statistique La théorie et ses applications** , Deuxième édition , Springer-Verlag France, Paris ;2010, p76.

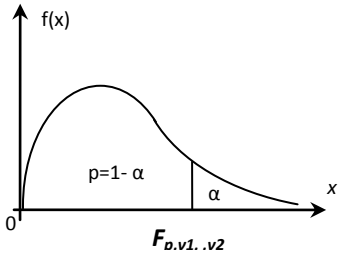
**ب. خصائص توزيع فيشر:**

$$\mu = \frac{\nu_2}{\nu_2 - 2} \quad (\nu_2 > 2) \quad , \quad \sigma^2 = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)(\nu_2 - 2)^2} \quad (\nu_2 > 4)$$

ويظهر من المعادلة تبعية منحنى f(x) بالإضافة ل x إلى كل من $\nu_1$ و$\nu_2$ ولذلك تحدد أي نقطة F من خلال ثلاثة معالم: $\nu_1$ و$\nu_2$ ،p (المساحة تحت المنحنى على يسار النقطة F) ، ونكتب $F_{p, \nu_1, \nu_2}$

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

**الشكل.10: تعيين قيمة F يتم في الجدول يتم من خلال $\nu_1, \nu_2$ و P**




**المصدر:** بوعبد الله صالح؛ **محاضرات الإحصاء الرياضي**، كلية العلوم الاقتصادية، جامعة المسيلة، الجزائر، السنة الجامعية 2005، 2006، ص.72

**نظرية 1**.

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

**نظرية 2**.

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

**نظرية 3**.

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

**مثال.18 :**

اوجد القيم التالية : $F_{(0.05;1;7)}$ $F_{(0.025;3;5)}$ $F_{(0.01;12;8)}$

**الحل:**

لايجاد $F_{(0.05;1;7)}$ نختار من جداول التوزيع، التوزيع الخاص بقيمة $\alpha=0.05$ ومن الصف في اعلى الجدول الخاص بدرجات حرية البسط نختار القيمة 1 حيث يوجد أسفلها عمود من الاعداد، ومن العمود أقصى اليسار والخاص بدرجة حرية المقام نختار القيمة 7 حيث يوجد على يمينه صف من الاعداد، وبتقاطع هذا الصف مع العمود أسفل درجة حرية البسط نجد أن:

$F_{(0.05;1;7)} = 5.59$

وبالمثل يكون:

$F_{(0.025;3;5)} = 7.76$

$F_{(0.01;12;8)} = 2.50$

## 3.1. ملخص مختصر لمحور المفاهيم الأساسية للمتغيرة العشوائية وأهم التوزيعات الاحتمالية:

يمكن تلخيص أهم ما تضمنه هذا المحور في الجدول التالي:

**الجدول.01: ملخص مختصر لأهم التوزيعات الاحتمالية المستمرة**

| التوزيع | المتغيرة العشوائية | أهم ما يجب معرفته عن دالة الكثافة |
|---|---|---|
| **التوزيع الطبيعي المعياري** $X \sim N(0, 1)$ | $f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \qquad -\infty < z < \infty$ مع : $E(Z) = 0,\ V(Z) = 1$ | $P(Z \le -z) = 1 - P(Z \le z) = P(Z \ge z)$<br>$P(-\sigma \le X \le \sigma) = P(-1 \le Z \le 1) = 0.6837$<br>$P(-2\sigma \le X \le 2\sigma) = P(-2 \le Z \le 2) = 0.9544$<br>$P(-3\sigma \le X \le 3\sigma) = P(-3 \le Z \le 3) = 0.9973$ |
| **توزيع ك2** $X \sim \chi^2_v$ | إذا كانت Xi متغيرات عشوائية مستقلة كل منها تتبع التوزيع الطبيعي المعياري، و $X = X_1^2 + X_2^2 + \ldots + X_v^2$ إذن: $X \sim \chi^2_v$ | $f(x) = 0 \quad si \quad x \le 0$<br>$E(X) = v, \quad V(X) = 2v$ |
| **توزيع ستيودنت** $T \sim t_v$ | لتكن المتغيرتان العشوائيتان المستقلتان Y وZ حيث: $Y \sim N(0, 1)$ و $Z \sim \chi^2_v$ ؛ إذن: $T = \frac{Y}{\sqrt{Z/v}} \sim t_v$ | $E(T) = 0,$<br>$V(T) = v/(v-2) \quad si \quad (v > 2)$ |
| **توزيع فيشر** $X \sim F_{v1, v2}$ | إذا كانت لدينا متغيرتان عشوائيتان مستقلتان حيث: $X_1 \sim \chi^2_{v_1}$ و $X_2 \sim \chi^2_{v_2}$ ، فإن $X = \frac{X_1/v_1}{X_2/v_2} \sim F_{v1, v2}$ | $f(x) = 0 \quad si \quad x \le 0$<br>$F_{p,v,\infty} = \frac{\chi^2_{p,v}}{v}$<br>$F_{1-p,\,1,\,v} = t^2_{1-(p/2),v}$<br>$F_{1-p,v_1,v_2} = 1/F_{p,v_2,v_1}$ |

المصدر: من اعداد الباحث اعتمادا على النظريات السابقة

# المحاضرة الثالثة:

## نظرية المعاينة-الجزء الأول

## 2. نظرية المعاينة

**تمهيد:**

يعتمد أسلوب المعاينة على تقدير المعالم الرئيسية للمجتمع من خالل بيانات أخذت من عينة ممثلة للمجتمع ما أي تتوافر فيها خصائص المجتمع الأصلي، وهذا لتخفيض أخطاء المعاينة إلى حدها الأدنى ويرتبط تمثيل العينة في المجتمع بعوامل عديدة كحجم العينة، تباين وخصائص المجتمع وكذا طريقة اختيار العينة، وكل هذه العناصر سيتم التطرق اليها في المحور الثاني من هذه المطبوعة حيث في بدايته سنتناول المفاهيم المتعلقة بنظرية المعاينة ثم نعرض مختلف الخصائص الإحصائية للمجتمع والعينة.

### 1.2. مفاهيم أساسية حول نظرية المعاينة

**أولا: المجتمع الاحصائي**

عبارة عن جميع القيم أو المفردات التي يمكن ان يأخذها المتغير ، فلو كانت دراستنا حول اطوال طلبة جامعة ما فان المجتمع في هذه الحالة هو أطوال جميع الطلبة في تلك الجامعة ، والمجتمع اما أن يكون مجتمعا محدودا أي ممكن حصر عدد مفرداتها ، أو يكون مجتمعا غير محدودا وهو المجتمع الذي من الصعب حصر عدد مفرداته مثل مجتمع نوع سمك معين في نهر معين أو عدد البكتريا في حقل ، مما سبق نالحظ أن المجتمع يقصد به القياسات أو القيم أو االشياء التي تم قياسها.

**ثانيا: العينة:**

يطلق على عملية اختيار العينة بالمعاينة وليست الفكرة في مجرد اختيار جزء من المجتمع فقط أيا كان و لكن من الضروري أن تتوفر لدينا بعض المعلومات عن المجتمع و التي تسمح لنا بالاختيار بعض مفرداته على اسس معينة تستطيع من خالل دراستها معرفة كثير عن هذا المجتمع ، أو يمكن تعريف العينة على أنها ذلك الجزء من المجتمع فالعينة عبارة عن مجموعة من المشاهدات أختيرت بطريقة ما من المجتمع ففي بعض االحيان دراسة المجتمع ككل قد يكون صعبا أو يحتاج الى وقت وجهد ومال فيستعاض عن دراسة المجتمع بدراسة العينة وصفاتها ومنها يستنتج خواص المجتمع االصلي الذي اخذت منه العينة ونرمز لمفردات المجتمع بـ N ومفردات العينة بـ n
فمصطلح المجتمع هنا قد 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، أما العينة فهي عادة ما تكون محدودة (علوان مطلق، 2009، ص.19).

### أ. العينة العشوائية Echantillon aléatoire

من أجل أن تكون العينة ممثلة للمجتمع، أحد الطرق المستخدمة هي العينة العشوائية. نظريا (قد يصعب تحقيق ذلك في الواقع)، نقول عن عينة أنها عشوائية إذا كان لكل مفردة في المجتمع نفس الاحتمال لأن تكون في العينة. تسمى هذه العينة بالعينة العشوائية 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 العشوائية (علوان مطلق، 2009، ص.31).
حيث تنقسم العينة الى عينة نفادية وأخرى غير نفادية وتفصيل ذلك فيما يلي:

**ب. العينة النفادية والعينة غير النفادية Echantillon exhaustif et non exhaustif** (صالح، 2006، ص.80):

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، وبالتالي فالعينة النفاذية تكون في حالة السحب بدون ارجاع وتعطى بالعلاقة التالية:

$$C^{n}_{N}$$

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.

وبالتالي فالعينة غير النفاذية تكون في حالة السحب مع الارجاع وتعطى بالعلاقة التالية:

$$N^{n}$$

**ثالثا: معالم المجتمع Paramètre d'une population** (موراي، جون، و ألو، 2004 ، ص.70):

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 من التوزيعات الاخرى.

ولتقدير معالم المجتمع من متوسط وتباين فإننا ننطلق من بيانات العينة ، ونسمي كل قيمة تحسب انطالقا من هذه البيانات بإحصائية المعاينة ، ونظريا فان احصائية المعاينة هي كل دالة في المتغيرات العشوائية التي تمثل القيم المحصل عليها في العينة.

و سنرمز للمتوسط المجتمع بالرمز $\mu$ حيث:

$$\mu = \sum xi / N$$

أما تباين المجتمع سنرمز له $\sigma^2$ حيث:

$$\sigma^2 = \sum (xi - \mu)^2 / N$$

**رابعا: إحصائية المعاينة Statistique de l'échantillonnage** (موراي، جون، و ألو، 2004، ص.71):

لتقدير معالم المجتمع (متوسط المجتمع $\mu$، تباين المجتمع $\sigma^2$ النسبة p ...) ننطلق من بيانات العينة، حيث نحتاج إلى حساب معالم مثل متوسط العينة $\bar{X}$ ، 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 العشوائية التي تمثل القيم المحصل عليها في العينة.

### 2.2. توزيع المعاينة للمتوسطات

اذا اخذنا عينات متكررة من مجتمع ما وقمنا بقياس متوسط لكل عينة، فإننا نجد أن هذه المتوسطات تختلف عن بعضها البعض، ويسمى التوزيع الاحتمالي لمتوسطات العينات هذه بـ " **توزيع المعاينة للمتوسطات**" ، والذي بدوره له أيضا متوسط معبر عنه بالرمز $\mu_{\bar{x}i}$ وانحراف معياري أو خطأ معياري $\sigma_{\bar{x}i}$ ، حيث سيتم الاعتماد على هذه الرموز في مختلف محاور المطبوعة، فمثلا لو كان لدينا 10 عينات عشوائية مسحوبة من مجتمع ما وحجم كل عينة **n= 10** وتم حساب المتوسط الحسابي لكل عينة وهو متغير عشوائي والتوزيع التكراري لمتوسطات هذه العينات يسمى بتوزيع المعاينة للمتوسط الحسابي وبصفة عامة نجد أن توزيع المعاينة للمتوسط المكون من كل العينات الممكن اخذها له خصائص مهمة ومفيدة في دراسة المجتمع عن طريق المعاينة.

#### 1.2.2. الوسط الحسابي لتوزيع المعاينة للمتوسطات (صالح، 2006، ص.81):

**نظرية.1:**

اذا كان $\mu$ يمثل متوسط مجتمع ما، و$\mu_{\bar{x}i}$ يمثل الوسط الحسابي لعينة مسحوبة من ذات المجتمع، فأن القيمة المتوقعة لمتوسط العينة $E(\mu_{\bar{x}i})$ يعبر عنها كما يأتي:

$$E(\mu_{\bar{x}i}) = \mu_{\bar{x}i}$$

ولإثبات ذلك نتبع الخطوات التالية:

$$E(\mu_{\bar{x}i}) = E(1/n \sum xi) = 1/n \sum E(xi) = 1/n \sum \mu = 1/n\, \mu n = \mu$$

#### 2.2.2. تباين توزيع المعاينة للمتوسطات :(جلال و مصطفى، 1990، ص.55)

**نظرية.2:**

اذا كان $\mu$ يمثل متوسط مجتمع ما $\mu_{\bar{x}i}$ يمثل متوسط عينة مسحوبة من ذات المجتمع بالإرجاع، فأن التباين $\sigma_{\bar{x}i}^2$ تباين توزيع المعاينة للمتوسطات يكتب كالتالي:

- ✓ **في حالة كان السحب بالارجاع**

$$\sigma_{\bar{x}i}^2 = \frac{\sigma^2}{n}$$

حيث n يمثل حجم العينة.

- ✓ **أما في حالة كان السحب بدون ارجاع**

$$\sigma_{\bar{X}i}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$$

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

**مثال.01:**

إذا كان لدينا مجتمع يمثل أعمار خمسة أطفال، وكانت أعمارهم على التوالي: 1، 3، 5، 6، 8.

**المطلوب :**

- أحسب الوسط الحسابي $\mu$ للمجتمع ؟
- بفرض أننا سحبنا كل العينات العشوائية الممكنة من حجم طفلين من هذا المجتمع باعتبار أن عملية السحب تمت بالارجاع ، ثم بدون ارجاع.
- احسب الوسط الحسابي لتوزيع معاينة الوسط الحسابي؟

**الحل:**

العينات الممكنة ذات الحجم n = 2 من مجتمع حجمه 5 مفردات ياعتبار السحب بالارجاع تحسب وفق العلاقة التالية:

$$N^n = 5^2 = 25$$

أي أنه هناك امكانية للحصول على 25 عينة مكونة من مفردتين من مجتمع مكون من 5 مفردات باعتبار انها مسحوبة **بالإرجاع** عناصرها ممثلة في الجدول أدناه، (الحالة 01)

العينات الممكنة ذات الحجم n = 2 من مجتمع حجمه 5 مفردات ياعتبار السحب تم بدون ارجاع تحسب وفق العلاقة التالية:

$$C^n{}_N = C_5{}^2 = 10$$

أي أنه هناك امكانية للحصول على 10 عينات مكونة من مفردتين من مجتمع مكون من 5 مفردات باعتبار انها مسحوبة **بدون ارجاع** عناصرها ممثلة في الجدول أدناه، (الحالة 02)

| **الحالة رقم 01، حالة السحب بالارجاع** | | **الحالة رقم 02، حالة السحب بدون ارجاع** | |
|---|---|---|---|
| **العينات الممكنة في حالة السحب بالارجاع (معاينة غير نفادية)** | **المتوسطات الممكنة $\bar{X}i$** | **العينات الممكنة في حالة السحب بدون ارجاع (معاينة نفادية)** | **المتوسطات الممكنة $\bar{X}i$** |
| (1, 1) | 1 | (1, 3) | 2 |
| (1, 3) | 2 | (1, 5) | 3 |
| (1, 5) | 3 | (1, 6) | 3.5 |
| (1, 6) | 3.5 | (1, 8) | 4.5 |
| (1, 8) | 4.5 | (3, 5) | 4 |
| (3, 1) | 2 | (3, 6) | 4.5 |
| (3, 3) | 3 | (3, 8) | 5.5 |
| (3, 5) | 4 | (5, 6) | 5.5 |
| (3, 6) | 4.5 | (5, 8) | 6.5 |
| (3, 8) | 5.5 | (8, 6) | 7 |
| (5, 1) | 3 | | |
| (5, 3) | 4 | | |
| (5, 5) | 5 | | |
| (5, 6) | 5.5 | | |
| (5, 8) | 6.5 | | |
| (6, 1) | 3.5 | | |
| (6, 3) | 4.5 | | |
| (6, 5) | 5.5 | | |
| (6, 6) | 6 | | |
| (6, 8) | 7 | | |
| (8, 1) | 4.5 | | |
| (8, 3) | 5.5 | | |
| (8, 5) | 6.5 | | |
| (8, 6) | 7 | | |
| (8, 8) | 8 | | |

أ. حساب الوسط الحسابي للمجتمع:

$$\mu = (\sum X_i) / N \Longrightarrow \mu = (1 + 3 + 5 + 6 + 8)/5 = 4.6$$

ب. حساب الوسط الحسابي لتوزيع المعاينة للوسط الحسابي في الحالتين عند السحب بالارجاع وبدون ارجاع.

ب.1. حالة السحب بالارجاع

$\mu_{\bar{x}i} = (\sum \bar{X}_i) / N^n$

$\mu_{\bar{x}i}=(1+2+3+3.5+4.5+2+3+4+4.5+5.5+3+4+5+5.5+6.5+3.5+4.5+5.5+6+7+4.5+5.5+6.5+7+8) / 25$

$\mu_{\bar{x}i} = 4.6$

ب.2. حالة السحب بدون ارجاع

$\mu_{\bar{x}i} = (\sum \bar{X}_i) / C^n_N$

$\mu_{\bar{x}i}=(2+3+3.5+4.5+4+4.5+5.5+5.5+6.5+7) / 10$

$\mu_{\bar{x}i} = 4.6$

**ملاحظة :**

من خلال حسابنا للوسط الحسابي للمجتمع والوسط الحسابي لمجموع متوسطات المعاينة في الحالتين في حالة تم السحب بالارجاع أو بدون ارجاع توصلنا الى أن: $\mu = \mu_{\bar{x}i}$ وهو ما نصت عليه **النظرية .01**

**مثال.02:**

أحسب تباين المجتمع في المثال.01، أحسب التباين (والانحراف المعياري) لتوزيع المعاينة للمتوسطات $\sigma^2_{\bar{x}i}$ علما أن العينة مسحوبة بالإرجاع (غير نفادية)، قارن بين تباين المجتمع وتباين متوسطات العينات الممكنة (توزيع المعاينة للمتوسطات).

**الحل:**

- حساب تباين توزيع المعاينة للمتوسطات

**أ.حالة المعاينة بالإرجاع**

**لدينا:**

$\sigma^2_{\bar{X}i} = [\sum_i (\bar{X}_i - \mu_{\bar{x}i})^2 ]/25 = 2.92;$

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

المقارنة:

**2.92 = 5.84 / 2**

أي أن:

$$\sigma^2_{\bar{x}i} = \frac{\sigma^2}{n}$$

هذا المثال يمهد للنظرية التالية:

**نظرية .2:**

إذا كانت X متغير عشوائي يمثل مجتمع ما و$\overline{X}_i$ متغيرة عشوائية تمثل متوسط عينة مسحوبة من ذات المجتمع بالإرجاع، فإن تباين $\overline{X}_i$ (تباين توزيع المعاينة للمتوسطات) يكتب كما يلي:

$$\sigma^2_{\overline{x}i} = \frac{\sigma^2}{n}$$ حيث n حجم العينة.

**البرهان** (صالح، 2006، ص.82)**:**

لنرمز ب Xi لقيم المتغيرة الأصلية X .

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

**ب. حالة المعاينة بدون إرجاع.**

في **المثال.1** أحسب تباين المتوسطات الممكنة للعينة $\sigma^2_{\overline{X}i}$ في حالة المعاينة بدون إرجاع، قارن بين تباين المجتمع وتباين المتوسطات الممكنة للعينة.

**الحل:**

**لدينا:**

$\sigma^2_{\overline{X}i} = [\sum_i (\overline{X}_i - \mu_{\overline{x}i})^2 ]/10 = 2.19$

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

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 .3:**

إذا كانت X متغيرة عشوائية تمثل مجتمع ما حجمه N و$\overline{X}_i$ متغيرة ع تمثل متوسط عينة حجمها n مسحوبة من ذات المجتمع بدون إرجاع، فإن تباين $\overline{X}_i$ (تباين توزيع المعاينة للمتوسطات) يكتب كما يلي: $\sigma^2_{\overline{x}i} = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$

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

**3.2.2. طبيعة توزيع $\overline{X}_i$**

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

**نظرية.1:**

إذا كان لدينا مجتمع موزع توزيعا طبيعيا بمتوسط $\mu$ وتباين $\sigma^2$ فإن متوسط العينة المسحوبة منه يتبع أيضا التوزيع الطبيعي بمتوسط $\mu$ وتباين $\sigma^2/n$،

ونكتب: $\overline{X}_i \approx N(\mu, \sigma^2/n)$

**نظرية.2:**

إذا كان المجتمع الذي تسحب منه العينة ذو متوسط $\mu$ وتباين $\sigma^2$ لكن ليس بالضرورة موزع طبيعيا فإن المتغيرة المعيارية z لـ $\bar{x}_i$ حيث: $z = \frac{\bar{x}_i - \mu}{\sigma / \sqrt{n}}$ تؤول إلى التوزيع الطبيعي المعياري عندما يكون n كبيرا ($n \geq 30$) ونكتب: $z \approx N(0, 1)$ **منطوق هذه النظرية هو ما يعرف بنظرية النهاية المركزية** (جلال و مصطفى، 1990، ص.55).

أما في حالة كان المجتمع محدود والمعاينة نفادية فنستبدل العبارة $\sigma/\sqrt{n}$ بالعبارة:

$$\sigma_{\bar{x}i} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

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

**مثال.03:**

مجتمع حجمه 900 بمتوسط $\mu=20$ وانحراف معياري $\sigma = 12$ ، بعد استخراج جميع العينات الممكنة ذات الحجم ($n = 36$) والحجم ($n = 64$).

**المطلوب:**

أحسب المتوسط والانحراف المعياري لتوزيع المعاينة للمتوسطات في الحالتين:

أ. في الحالة (1): حجم العينة $n = 36$

ب. في الحالة (2): حجم العينة $n = 64$

**الحل:**

**حساب الوسط الحسابي لمجموع متوسطات المعاينة**

**نعلم أن:**

$$\mu_{\bar{x}i} = \mu$$

وبالتالي:

$$\mu_{\bar{x}i} = 20$$

**أ. في الحالة (1): حجم العينة n = 36**

أولا نقوم بحساب النسبة (n/N):

لدينا:

$n/N = 36/900 = 0.04 < 0.05 \Rightarrow \sigma_{\bar{x}i} = \sigma/\sqrt{n} = 12/\sqrt{36} = 2$

**ب. في الحالة (2): حجم العينة n = 64**

أولا نقوم بحساب النسبة (n/N):

لدينا:

$n/N = 64/900 = 0.071 >$

$0.05$

$\Rightarrow \sigma_{\bar{x}i} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$

$$\Rightarrow \sigma_{\bar{x}i} = \frac{12}{\sqrt{64}}\sqrt{\frac{900-64}{900-1}}$$

$$\Rightarrow \sigma_{\bar{X}i} = 1.92$$

**مثال.04:**

باستخدام معطيات المثال السابق باعتماد أن حجم العينة مساو لـ (n = 36) .

**المطلوب:**

- أحسب احتمال أن يكون $\bar{X}i$ محصورا بين 18 و22.
- أحسب نفس الاحتمال باعتماد أن حجم العينة مساو لـ (n = 64)

**الحل:**

**أ. في الحالة (1): حجم العينة n = 36**

لدينا:

$$P(18< \bar{X}<22) = P(Z_1< Z<Z_2)$$

وبالتالي يجب علينا حساب كل من : $Z_2$ , $Z_1$

$$Z_1 = \frac{\bar{X}i_1 - \mu}{\sigma/\sqrt{n}} = \frac{18-20}{12/\sqrt{36}} = -1 \ , \quad Z_2 = 1 \ \Rightarrow P(18 < \bar{X} < 22) = P(Z_1 < Z < Z_2) = 0.6827$$

**أ. في الحالة (2): حجم العينة n = 64**

**لدينا:**

$$Z_1 = \frac{\bar{X}i_1 - \mu}{\sigma/\sqrt{n}}\left(\frac{N-n}{N-1}\right) = \frac{18-20}{1.92} = -1.04 \ , \quad Z_2 = 1.04 \ \Rightarrow P(18 < \bar{X} < 22) = P(-1.04 < Z < 1.04) = 0.70$$

### 4.2.2. ملخص مختصرلأهم خصائص توزيع المعاينة للمتوسطات.

**الجدول.02: ملخص مختصرلأهم خصائص توزيع المعاينة للمتوسطات.**

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

المصدر: من اعداد الباحث اعتمادا على النظريات السابقة

**واجب منزلي:**

إذا كان لدينا مجتمع يمثل أعمار ثلاثة أطفال، وكانت أعمارهم على التوالي: 2، 4 ،6

المطلوب :أ-احسب الوسط الحسابي والانحراف المعياري للمجتمع ؟

ب- بفرض أننا سحبنا كل العينات العشوائية الممكنة من حجم طفلين من هذا المجتمع،

- ✓ أحسب الوسط الحسابي والانحراف المعياري لتوزيع معاينة للوسط الحسابي؟
- ✓ أحسب احتمال أن يكون عمر الطفل محصورا بين 2 و5.

### 3.2. توزيع المعاينة للنسبة

نفترض انه لدينا مجتمع غير محدود موزعا توزيعا ذي الحدين باحتمال وجود صفة معينة نرمز لها بالرمز (p) واحتمال عدم وجدو هذه الصفة بالرمز (q=1 -p)، وكمثال على ذلك يكون المجتمع هو كل رميات قطعة عملة متكاملة التوازن ويكون احتمال الحصول على الصورة 1/2 =p، لنعتبر أن كل العينات الممكنة ذات الحجم n المسحوبة من هذا المجتمع وفي كل عينة نحدد احصاء نسبة النجاح p وفي هذا المثال تكون النسبة تمثل نسبة عدد الصور التي نحصل عليها الى حجم العينة n وبالتالي نحصل على توزيع المعاينة الذي متوسطه $\mu_p$ وانحرافه المعياري $\sigma_p$ في الصورة (موراي، جون، و ألو، 2004، ص.77).

حيث : $\dot{p}=na/n$ $P=Na/N$

**مع:**

Na:عدد المفردات التي تحقق خاصية ما في مجتمع

N : حجم المجتمع

na : عدد المفردات التي تحقق نفس الخاصية في حجم العينة

n : حجم العينة

**نظرية.01:**

لتكن المتغيرة العشوائية $X_i$ التي تمثل مجتمع ما غير محدود وموزع توزيعا طبيعيا حيث p نسبة المفردات في المجتمع ذات صفة معينة، ولتكن $\dot{p}$ متغيرة عشوائية تمثل نسبة المفردات ذات الصفة المذكورة في العينة المسحوبة من ذات المجتمع (نفس الصفة)، نحصل على توزيع للإحصائية $\dot{p}$ حيث معالمه: $E(\dot{p})$ و$\sigma_{\dot{p}}$، هذه المعالم تساوي (موراي، جون، و ألو، 2004، ص.77):

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

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

**أي أنه** اذا كانت P تمثل النسبة في المجتمع تحقق خاصية ما و$\dot{p}$ تمثل النسبة في العينة التي تحقق نفس الخاصية فان توزيع المعاينة يقترب من التوزيع الطبيعي حسب نظرية النهاية المركزية لما $(n \geq 30)$ مع $n(1-p) > 5$ .

**ملاحظة:**

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

**مثال.01:**

يحتوي احدى المستشفيات على 800 مريض من بينهم 320 مريض يعاني من مرض مزمن، اذا سحبنا عينة من 40 مريض بدون ارجاع.

**المطلوب:**

- حدد طبيعة توزيع المرضى المصابين بمرض مزمن في العينة
- ما هو احتمال ان تكون نسبة المصابين بمرض مزمن في العينة أقل من 50%

**الحل:**

$P = Na/N = 320/800 = 0.4$ N=800 n=40

نلاحظ ان حجم العينة $(n = 40 >= 30)$ وعليه فان توزيع المعاينة النسبة ṗ يقترب من التوزيع الطبيعي حسب نظرية النهاية المركزية مع $n(1-p) = 40\ (1-0.4) = 24 > 5$ .

وبما اننا في حالة السحب بدون ارجاع يصبح لدينا:

$$\dot{p} \approx N\ (p, \sigma_{\dot{p}})$$

$$\dot{p} \approx N\ (p, pq/n*(N-n/N-1))$$

**مع:**

$$\mu_{\dot{p}} = p = 0.4$$

$$\sigma_{\dot{p}} = \sqrt{\frac{P(1-p)}{n}}\sqrt{\left(\frac{N-n}{N-1}\right)} = \sqrt{\frac{0.4(1-0.4)}{40}}\sqrt{\left(\frac{800-40}{800-1}\right)} = 0.75$$

**اذا :**

$$\dot{p} \approx N\ (0.4, 0.75)$$

حساب احتمال أن تكون نسبة المصابين بمرض مزمن في العينة أقل من50% :

$$P(\dot{p}<0.5) = P(Z<0.5-0.4/0.75) = P(Z<0.13) = 0.5517$$

**مثال.02:**

لاحظت إدارة الجامعة أنه في عينة مكونة من 100 طالب، 40 طالب تحصلوا أخيرا على الشهادة الجامعية، تريد الإدارة تقدير نسبة الطلبة الذين يحصلون على الشهادة داخل مجال يكون احتماله 90 بالمائة.

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

نفترض أن N كبير بحيث : $n/N < 0.05$

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

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

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

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

# المحاضرة الثالثة:

## نظرية المعاينة-الجزء الثاني

### 4.2. توزيع المعاينة للفروق والمجاميع

#### 1.4.2. متوسط وتباين توزيع المعاينة للفروق والمجاميع

اذا كان لدينا مجتمعين حجمهما $N_1$ و$N_2$ نسحب من كل منهما عينة عشوائية حجمها على التوالي $n_1$ و$n_2$ ، نحسب في كل عينة مسحوبة من المجتمع الأول الإحصائية $S_1$ ونحسب نفس الاحصائية (المتوسط مثلا أو التباين ...) في كل عينة من المجتمع الثاني ونسميها $S_2$. إن الفرق بين الاحصائيتين $S_1 - S_2$ يشكل بدوره متغيرة عشوائية لها المتوسط والتباين التاليين (موراي، جون، و ألو، 2004، ص.ص، 78.79):

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

فإذا كانت الاحصائية هي **المتوسط** فإن:

$$\mu_{\bar{X}1-\bar{X}2} = \mu_{\bar{X}1} - \mu_{\bar{X}2} = \mu_1 - \mu_2$$
$$\sigma^2_{\bar{X}1-\bar{X}2} = \sigma^2_{\bar{X}1} + \sigma^2_{\bar{X}2} = \sigma^2/n_1 + \sigma^2/n_2$$

أما إذا كانت الاحصائية هي **النسبة** فإن:

$$\mu_{\dot{p}1-\dot{p}2} = \mu_{\dot{p}1} - \mu_{\dot{p}2} = p_1 - p_2$$
$$\sigma^2_{\dot{p}1-\dot{p}2} = \sigma^2_{\dot{p}1} + \sigma^2_{\dot{p}2} = p_1q_1/n_1 + p_2q_2/n_2$$

وإذا كان الاهتمام بمجموع الاحصائيتين بدلا من الفرق بينهما فإن:

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

#### 2.4.2. طبيعة توزيع المعاينة للفرق بين متوسطين

إذا كان المجتمعين الذين سحبت منه العينتين ليس بالضرورة موزعين توزيعا طبيعيا وكانت العينتين المسحوبتين حجمهما $n_1 \geq 30$ و $n_2 \geq 30$ ، فبالاستدلال بمنطوق نظرية النهاية المركزية فان توزيع المتغيرة المعيارية للفرق بين متوسطين يقترب من التوزيع الطبيعي المعياري. ونكتب (صالح، 2006، ص.85):

$$\mu_{\bar{x}1-\bar{x}2} \approx N(0, 1)$$

**مثال.01:**

اذا كان لدينا مجتمعين الأول وسطه الحسابي يساوي 30 وتباينه 25، والثاني وسطه الحسابي يساوي 20 وتباينه 16، وسحبت منهما عينتين حجمها على التوالي 30 مشاهدة و35 مشاهدة.

**المطلوب:**

- أوجد توزيع الفرق بين متوسطي العينتين؛
- احسب الاحتمال $P(\bar{X}_1 - \bar{X}_2 < 12)$

**الحل:**

**لدينا:**

$$\bar{X}_1 - \bar{X}_2 \approx N(\mu_{\bar{x}1 - \bar{x}2} ; \sigma^2_{\bar{X}1 - \bar{X}2})$$

**مما سبق يصبح لدينا:**

$$\bar{X}_1 - \bar{X}_2 \approx N(\mu_1 - \mu_2 ; \sigma^2/n_1 + \sigma^2/n_2)$$

**فنحصل على:**

$$\bar{X}_1 - \bar{X}_2 \approx N(30 - 20 ; 25/30 + 16/35)$$

$$\bar{X}_1 - \bar{X}_2 \approx N(10 ; 1.29)$$

**وبالتالي:**

$$P(\bar{X}_1 - \bar{X}_2 < 12) = P(Z<(\bar{X}_1 - \bar{X}_2) - 12 / \sqrt{1.29}) = P(Z<1.76) = 0.9608$$

**مثال.02**:

اذا كانت نسبة النجاح في امتحان توجهي في مدرسة للبنات P هي 7.0 ؛ وكانت نسبة النجاح في ذلك الامتحان في مدرسة للذكور هي 0.65 واخذت عينة عشوائية حجمها $n_1=70$ من مدرسة البنات وأخذت عينة عشوائية حجمها $n2=35$ من مدرسة الذكور.

**المطلوب:**

ما هو احتمال ان تزيد نسبة النجاح في مدرسة البنات عن نسبة النجاح في مدرسة الذكور بمقدار 0.10 على الاكثر؟

**الحل:**

حساب احتمال ان تزيد نسبة النجاح في مدرسة البنات عن نسبة النجاح في مدرسة الذكور بمقدار 0.10 على الاكثر، يعني حساب ما يلي:

$$P(\dot{p}1 - \dot{p}2 < 0.10) = ?$$

**لدينا:**

$$P(\dot{P}_1 - \dot{P}_2 < 0.10) = P\left(\frac{(\dot{P}1 - \dot{P}2) - (P1 - P2)}{\sqrt{\frac{P(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}} < \frac{0.10 - (0.70 - 0.65)}{\sqrt{\frac{0.7(1-0.7)}{70} + \frac{0.(1-0.65)}{35}}}\right)$$

$$= P\left(Z < \frac{0.05}{0.097}\right) = P(Z < 0.51)$$

$$= 0.6950$$

### 5.2. توزيع المعاينة للتباين وتوزيع المعاينة لنسبة تباييني عينتين

#### 1.5.2. توزيع المعاينة للتباين (موراي، جون، و ألو، 2004، ص.ص، 81.80):

**نظرية.01:**

اذا كانت $x_i$ متغيرة عشوائية تمثل مجتمع ما و$S^2$ متغيرة عشوائية تمثل تباين عينة مسحوبة بالإرجاع (أو بدون إرجاع من مجتمع غير محدود) حجمها n، فان : $E(S^2) = \mu_{S^2} = \sigma^2\left(\frac{n-1}{n}\right)$

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

**البرهان** (صالح، 2006، ص.86):

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

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

**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" لـ $\sigma^2$ ويرمز له بـ: $\hat{S}^2$ حيث:

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

**نظرية.02** (صالح، 2006، ص.86):

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.01:**

ليكن لدينا مجتمع طبيعي حجمه 100 نسحب منه عينة حجمها $n = 16$ . ما هو احتمال أن يكون تباين العينة $S^2$ أقل من أو يساوي 10 علما أن تباين المجتمع 80.

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

من الجدول $P(X^2_{15} \leq 2) < 0.005$

**نظرية.03** (صالح، 2006، ص.87):

إذا كانت م ع تمثل مجتمع ما محدود و$S^2$ متغيرة 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 1)

**2.5.2. توزيع المعاينة لنسبة تباينين**

رأينا فيما سبق أن: $X = \frac{X_1/\nu_1}{X_2/\nu_2} \sim F_{\nu1,\nu2}$ في حالة المتغيرتان العشوائيتان مستقلتان مع $X_1 \sim \chi^2_{\nu_1}$ و $X_2 \sim \chi^2_{\nu_2}$ نستنتج ما يلي:

**نظرية.04** (صالح، 2006، ص.87):

ليكن لدينا مجتمعان طبيعيان تبايناهما $\sigma^2_1$ , $\sigma^2_2$ . وسحب منهما عينتين عشوائيتين حجمهما على التوالي: $n_1$ , $n_2$ فان:

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

**مثال.02:**

عينتين حجمهما 8 و10 مسحوبتين من مجتمعين طبيعيين تباينهما على التوالي 20 و36. ما احتمال أن يكون تباين الأولى أكبر من ضعف تباين الثانية؟

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

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

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

**3.5.2. تذكير**

**أ. الانحراف المعياري لتوزيع المعاينة للتباين**

**نظرية.05** (صالح، 2006، ص.88)**:** إذا كانت X م ع تمثل مجتمع ما و$S^2$ متغيرة ع تمثل تباين عينة مسحوبة من ذات المجتمع، فإن:

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

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

**ب. الانحراف المعياري لتوزيع المعاينة للانحراف المعياري**

$$\sigma_S = \begin{cases} \frac{\sigma}{\sqrt{2/n}} & \text{si } X \sim N \quad \text{ou} \quad X \approx N \\ \sqrt{\frac{\mu^4 - \sigma^4}{4n\sigma^2}} & \text{sinon} \end{cases}$$

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

**الجدول أدناه يلخص جميع ما ورد في النظريات السابقة:**

**الجدول.03: ملخص نظريات توزيع المعاينة للنسبة، للفرق، للتباين ولنسبة تباينين**

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

المصدر: من اعداد الباحث اعتمادا على النظريات السابقة

# المحاضرة الرابعة:

## نظرية التقدير-الجزء الأول

## 3. نظرية التقدير

تطرقنا في المحور السابق 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،

فالتقدير يمثل لنا تلك الأداة التي تسمح لنا بإجراء استدلالات حول معالم المجتمع انطلاقا من عينة مسحوبة منه، هذا الاستدلال لا يصح إلا في ظل شروط محددة، تتعلق أساسا بفرضيات تتمحور حول طبيعة بيانات المجتمع والعينة لذا سنتطرق في هذا المحور الى أهم المفاهيم الاساسية المتعلقة بالتقدير ثم نتناول مختلف مجالات الثقة للمعالم الاحصائية، حيث سنتناول كل هذا بالتفصيل

فيما يلي:

### 1.3. مفاهيم أساسية

#### 1.1.3. مفهوم التقدير (السقاف، 2020، ص.115):

المقصود بالتقدير Estimation تقدير معالم المجتمع أو (التوزيع الاحتمالي) والتي غالبا ما تكون مجهولة ويكون المطلوب هو الحصول على تقديرات لها من خلال معطيات العينة، ومثال ذلك تقدير متوسط دخل السكان في رقعة جغرافية ما أو تقدير متوسط عمر فئة معينة من المجتمع...الخ، عموما نميز بين نوعين من التقدير، الاول يسمى بالتقدير النقطي والثاني يسمى بالتقدير بمجال. وتفصيل هذا فيما يلي (السقاف، 2020، ص.ص، 116.115):

**ا-حالة التقدير النقطي:**

نعتمد على هذا النوع من التقدير عندما نحصل على قيمة واحدة من العينة كاحصاءة الوسط الحسابي وتستخدم هذه القيمة الواحدة كتقريب أو كتقدير للوسط الحسابي للمجتمع التي تكون مجهولة القيمة ومحل بحث من طرف الباحث، فمثلا لو أخذنا متوسط الدخل الشهري لعينة من الأسر في مدينة ما، فنكون قد تحصلنا على تقدير بنقطة لمتوسط دخل جميع الأسر في هذه المدينة.

**ب-حالة التقدير بمجال:**

أما في حالة التقدير وفق مجال فنتحصل على مجال معرف بحدين (حد أدنى وحد أعلى)، حيث نتحصل عليهما من بيانات العينة المسحوبة فمثالا إذا قدرنا الوسط الحسابي لأعمار الناخبين يتراوح ما بين 34 و46 سنة أي يتراوح بين 34 سنة كحد أدنى و46 سنة كحد أعلى، هنا نكون قد حصلنا على تقدير وفق مجال للوسط الحسابي لأعمار الناخبين في المجتمع ويحتوي المجال على أكثر من قيمة بل قد يكون غير نهائي وغير محدود غالبا، ويسمى التقدير بمجال الثقة ايضا لأن هذه الفترات تعتمد في تكوينها الاحصائي على درجات ثقة أو مستويات ثقة معينة مثل 95% أو 99%.

#### 2.1.3. بعض خصائص المقدر:

لتقدير معلمة من معالم مجتمع محل دراسة، نحتاج إلى اختيار الإحصائية المناسبة في العينة لتقدير هذه المعلمة؛ وغالبا ما تكون المعلمة المناظرة في العينة هي أحسن مقدر، كأن نقدر متوسط المجتمع $\mu$ من خلال متوسط العينة $\mu_{\bar{x}i}$. تسمى الإحصائية المستخدمة في التقدير **بالمقدر**؛ عموما هناك عديد الخصائص للحكم على جودة الإحصائية في العينة المناظرة للمعلمة:

**أ. المقدر غير المتحيز**

نقول عن إحصائية ما بأنها مقدر غير متحيز Sans biais لمعلمة المجتمع إذا كان متوسطها أو توقعها الرياضي مساويا لمعلمة المجتمع، أي أن مقدار الخطأ في التقدير يساوي صفرا وبعبارة رياضية نجد أن شرط عدم الانحياز هو $E(\hat{\theta}) = \theta$ ، حيث يعتبر شرط عدم التحيز أهم شروط التقديرات الوسيطية (السقاف، 2020، ص.116).

**مثال**:

نقول عن متوسط العينة $\bar{X}$ أنه مقدر غير متحيز لمتوسط المجتمع $\mu$ لأن $E(\bar{X}) = \mu$ . في المقابل نسمي الإحصائية $S^2$ في معاينة بالإرجاع أنها مقدر متحيز لـ $\sigma^2$ لأن:

$E(S^2) = \sigma^2 (n-1)/n \neq \sigma^2$ ، بينما تعتبر الاحصائية $\hat{S}^2 = S^2 n/(n-1)$ مقدرا غير متحيز في معاينة بالإرجاع.

**ب. الكفاءة**

تتعلق كفاءة (Efficacité) مقدر ما بمقدار التباين لتوزيع المعاينة للإحصائية، فإذا كان لمقدرين (إحصائيتين) نفس المتوسط نقول عن المقدر ذو توزيع المعاينة الأقل تباينا أنه الأكثر كفاءة، أي أنه اذا كان لدينا أكثر من مقدر لمعلمة معينة فنقول أن المقدر الأكفأ هو المقدر الذي له أقل تباين (تومي، 2011، ص.ص، 24.22).

**مثال**:

لكل من توزيعي المعاينة للمتوسط والوسيط نفس المتوسط هو متوسط المجتمع $\mu$ ، لكن يعتبر المتوسط $\bar{X}$ مقدرا أكثر كفاءة لمتوسط المجتمع $\mu$ من الوسيط لأن تباين توزيع المعاينة للمتوسطات $V(\bar{X}) = \sigma^2/n$ 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.

**ج. التقارب convergeance**

نقول عن مقدر أنه متقارب إذا كان يؤول إلى قيمة المعلمة المقدرة عندما يؤول حجم العينة إلى ما لا نهاية (السقاف، 2020، ص.116).

**مثال**: يعتبر متوسط العينة مقدرا متقاربا لمتوسط المجتمع لأن (السقاف، 2020، ص.116):

$$E(\bar{X}) = \mu \quad , \quad V(\bar{X}) = \frac{\sigma^2}{n} \xrightarrow[n \longrightarrow \infty]{} 0.$$

## 3.1.3. طرق التقدير

**أ. التقدير النقطي والتقدير بمجال** (صالح، 2006، ص.92).

قد نحتاج إلى تقدير لمعلمة مجتمع بقيمة واحدة ونقول عن هذا التقدير أنه **تقدير نقطي**، وأحيانا أخرى نحتاج إلى تقدير معلمة المجتمع بنقطتين يحددان مجال لقيمة المعلمة ونقول عن هذا النوع من التقدير أنه **تقدير بمجال**.

**مثال.01** :

إذا قدرنا دخل الأسرة في منطقة ما بـ 18000 دج، نكون قد قدرنا دخل الأسرة تقديرا نقطيا. ويكون تقديرنا بمجال إذا قلنا مثلا أن الدخل يساوي 18000 ± 2000 أي أنه يتراوح بين 16000 و20000دج.

**ب. درجة التأكد**

لكي يكون التقدير علميا ينبغي تقييم احتمال أن تكون المعلمة تنتمي فعلا إلى المجال المحدد، لذلك نلحق بالمجال ما يسمى بدرجة أو مستوى الثقة، ويرمز له بـ p. والاحتمال المعاكس يسمى باحتمال الخطأ ويرمز له ب α ، ويسمى أيضا بـ "مستوى المعنوية".

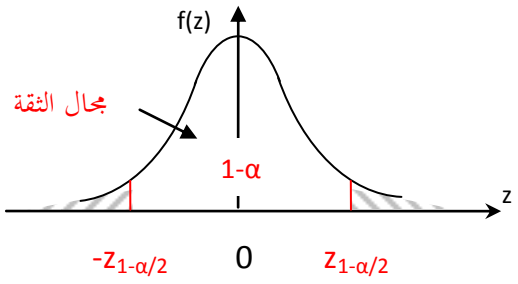
**مثال.02:**

دخل الأسرة في المنطقة (أ) ينتمي إلى المجال [16000، 20000] بمستوى معنوية 5 % أي بمستوى ثقة 95 % . وتسمي الحدود 16000 و20000 **بحدود الثقة**.

**ج. تعيين حدود مجال الثقة**

تحدد حدود الثقة من خلال معاملات الثقة التي بدورها تحدد من خلال مستوى المعنوية (مستوى الثقة). ففي حالة استخدام التوزيع الطبيعي للتقدير تكون القيمتين ± 1.96 معاملات الثقة من أجل مستوى ثقة 95% بينما القيمتين ±2.58 تمثلان معاملات الثقة من أجل مستوى ثقة 99 % .

**الشكل.11: مجال الثقة للتوزيع الطبيعي**




**المصدر:** بوعبد الله صالح؛ **محاضرات الإحصاء الرياضي**، كلية العلوم الاقتصادية، جامعة المسيلة، الجزائر، السنة الجامعية 2005، 2006، ص.92

**مثال.03:**

ليكن $\mu_s$ و$\sigma_s$ متوسط وانحراف معياري توزيع المعاينة لإحصائية ما s حيث $\mu_s = \mu$ . إذا كان توزيع المعاينة لـ s توزيعا طبيعيا (كما هو الحال بالنسبة لأغلب الإحصائيات عندما $(n \geq 30)$ ) فإننا نقدر مثلا وبالنظر إلى توزيع s أن:

القيمتين $\mu_s \pm 1.96\sigma_s$ تمثلان **حدود الثقة** بـ 95 %، و $\mu_s \pm 2.58\sigma_s$ حدود الثقة بـ 99 %.

في حالة التوزيع الطبيعي يرمز لحدود الثقة ب $Z_c$ أو $Z_{1-\alpha/2}$ (أنظر الرسم أعلاه).

### 2.3. فترة الثقة للمتوسط

#### 1.2.3. حالة المجتمع يتوزع توزيعا طبيعيا بتباين معلوم:

نستخدم التوزيع الطبيعي لتحديد مجال الثقة في هذه الحالة إذا علمنا أن المجتمع الذي سحبت منه العينة يتبع التوزيع الطبيعي، و/أو في حالة العينة الممتدة ($n \geq 30$) يمكن كذلك الاستفادة من نظرية النهاية المركزية (التي تخص في الحقيقة توزيع مجموع قيم العينة -في حالة كون العينة كبيرة بما فيه الكفاية- وليس المتوسط.) أن $\bar{X}$ تتبع التوزيع الطبيعي.

وبالتالي تكتب حدود مجال الثقة كما يلي (السقاف، 2020، ص.119):

$$\mathbf{IC\mu} = \bar{X} \pm z_c \cdot \frac{\sigma}{\sqrt{n}}$$

ونستخدم الصيغة أدناه إذا كان المجتمع محدود (ذا حجم N) والمعاينة نفادية كالآتي:

$$\mathbf{IC\mu} = \bar{X} \pm z_c \cdot \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

#### 2.2.3. حالة المجتمع يتوزع توزيعا طبيعيا بتباين مجهول (صالح، 2006، ص.93):

غالبا ما يكون الانحراف المعياري للمجتمع σ مجهولا، ولذلك نعوض σ في الصيغ السابقة بالمقدر S أو S'.

في حالة كان تباين المجتمع σ مجهول تكتب حدود مجال الثقة بالاستعانة باحصاءة تباين المعاينة وتكتب كما يلي:

$$\mathbf{IC\mu} = \bar{X} \pm z_c \cdot \frac{S}{\sqrt{n-1}} \quad \text{أو} \quad \mathbf{IC\mu} = \bar{X} \pm z_c \frac{S'}{\sqrt{n}}$$

الجدول الآتي يبين قيم $z_c$ التي تمثل حدود مجال الثقة بحسب مستوى الثقة :

**الجدول.04 : قيم $z_c$ لحدود مجال الثقة**

| **مستوى الثقة 1-α** | 0.99 | 0.98 | 0.95 | 0.90 | 0.8 | 0.5 |
|---|---|---|---|---|---|---|
| **α مستوى المعنوية** | 0.01 | 0.02 | 0.05 | 0.10 | 0.2 | 0.5 |
| **1- α/2** | 0.995 | 0.99 | 0.975 | 0.95 | 0.9 | 0.75 |
| **$Z_{1-\alpha/2}$** | 82.5 | 2.326 | 1.96 | 1.645 | 1.282 | 0.674 |

**المصدر:** من اعداد الباحث بالاعتماد على جدول التوزيع الطبيعي المعياري

**مثال.01** :

نقدر أن μ يوجد داخل المجال $\bar{X} \pm 1.96\sigma_{\bar{X}}$ بمستوى ثقة 95% (0.95) أي بمستوى معنوية 5 % (0.05)، وداخل المجال $\bar{X} \pm 2.58\sigma_{\bar{X}}$ بمستوى ثقة 99% أي بمستوى معنوية 0.01...الخ.

### 3.2.3. تقدير فترة الثقة للمتوسط باستخدام التوزيع t :

في حالة العينة الصغيرة (n < 30) وσ مجهول نستخدم توزيع ستيودنت لتحديد مجالات الثقة لـ μ. مثلا القيم $t_{0.975}$- ؛ $t_{0.975}$ تحد 95 % من المساحة تحت المنحنى ونقول أن $t_{0.975}$ ; $t_{0.975}$- تمثل القيم الحرجة أو معاملات الثقة عند مستوى ثقة 95 % ونكتب (السقاف، 2020، ص.122):

$$IC\mu = -t_{0.975} < \frac{\bar{X} - \mu}{\hat{S}/\sqrt{n}} < t_{0.975}$$

ومنه نستخلص مجال الثقة لـ μ كما يلي:

$$IC\mu = \bar{X} - t_{0.975} \cdot \frac{\hat{S}}{\sqrt{n}} < \mu < \bar{X} + t_{0.975} \cdot \frac{\hat{S}}{\sqrt{n}}$$

يمكننا تلخيص الحالات السابقة لبناء فترة الثقة من خلال الجدول الثالثي:

**الجدول.05: فترة الثقة حسب معلومية التباين والتوزيع الاحصائي**

| توزيع المجتمع | تباين المجتمع | حجم العينة n | الانحراف المعياري | الوسط الحسابي للمعاينة |
|---|---|---|---|---|
| المجتمع يتوزع توزيعا طبيعيا | معلوم | n≥30 ; n<30 | σ/√n | N~ (μ ; σ/√n) |
| | مجهول | n≥30 | S'/√n | N~ (μ ; S'/√n) |
| | | n<30 | S'/√n | $t_{\alpha; n-1}$ |
| المجتمع مجهول التوزيع الاحصائي | معلوم | n≥30 | σ/√n | N~ (μ ; σ/√n) |
| | مجهول | n≥100 | S'/√n | N~ (μ ; S'/√n) |

**المصدر:** من اعداد الباحث اعتمادا على العلاقات الواردة في النظريات السابقة

### 4.2.3. أمثلة:

**مثال.01:**

اذا كان لدينا مجتمع موزعا توزيعا طبيعيا، بانحراف معياري قدره 10و $\sigma =$ سحبت منه عينة عشوائية حجمها 24 متوسطها مساو لـ : $\bar{X} = 56$.

**المطلوب:**

بمستوى ثقة مساو لـ 95%، قم بتقدير الوسط الحسابي للمجتمع μ

**حل المثال.01:**

لدينا مستوى الثقة: 1-α = 95%

مستوى الخطأ او المعنوية : 5%

**ومنه: $Z_{1-\alpha/2} = 1.96$**

لتقدير الوسط الحسابي للمجتمع سنستعين بالعلاقة التالية:

$$\bar{X} - Z1-\alpha\,2 * \sigma/\sqrt{n} < \mu < \bar{X} + Z1-\alpha\,2 * \sigma/\sqrt{n}$$

**بالتعويض في العلاقة أعلاه نجد:**

$$56 - 1.96 * 10/\sqrt{24} < \mu < 56 + 1.96 * 10/\sqrt{24}$$

$$51.99916 < \mu < 60.00083$$

$$IC\mu = [56 \mp 10/\sqrt{24} * 1.96]$$

أي أنه عند مستوى ثقة مقدارها 95 % ستكون القيمة المقدرة للوسط الحسابي للمجتمع محصورة في المجال:

**[ 51.99916 ؛60.00083 ]**

**مثال.02:**

اذا كان لدينا مجتمع وسحبت منه عينة عشوائية حجمها 81 و متوسطها مساو لـ : $\bar{X} = 73$.

**المطلوب:**

اذا علمت أن قيمة الانحراف المعياري لتوزيع المعاينة للمتوسطات $\sigma_{\bar{x}i} = 6$ ، فقم بتقدير الوسط الحسابي للمجتمع $\mu$ عند مستوى ثقة مساو لـ 98%،

**حل المثال.2 :**

لدينا مستوى الثقة: $1-\alpha = 98\%$

مستوى الخطأ او المعنوية : 2%

**ومنه: $Z_{1-\alpha/2} = 2.23$**

كما يتضح لنا أن تباين المجتمع $\sigma^2$ مجهول، لكن حجم العينة اكبر من 30 ولذلك سنستعين بقيمة الانحراف المعياري لتوزيع المعاينة للمتوسطات $\sigma_{\bar{x}i}$ بدل الانحراف المعياري للمجتمع.

وبالتالي لتقدير الوسط الحسابي للمجتمع سنستعين بالعلاقة التالية:

$$\bar{X} - Z_{1-\alpha/2} * \sigma_{\bar{x}i} / \sqrt{n} < \mu < \bar{X} + Z_{1-\alpha/2} * \sigma_{\bar{x}i} / \sqrt{n}$$

**بالتعويض في العلاقة أعلاه نجد:**

$$73 - 2.33 * 6/\sqrt{81} < \mu < 73 + 2.33 * 6/\sqrt{81}$$

$$71.447 < \mu < 74.553$$

$$IC\mu = [73 \mp 6/\sqrt{81} * 2.33]$$

أي أنه عند مستوى ثقة مقدارها 98% ستكون القيمة المقدرة للوسط الحسابي للمجتمع محصورة في المجال:

**[ 71.447 ؛74.553 ]**

**مثال.03.**

اذا كانت دخول مجموعة من الأفراد في دولة ما تتبع التوزيع الطبيعي، وسحبت منه عينة عشوائية حجمها 10 أفراد، بمتوسط حسابي قدره $\bar{X} = 72$ وقيمة الانحراف المعياري لتوزيع المعاينة للمتوسطات $\sigma_{\bar{x}i} = 6.4$ ،

**المطلوب:**

بمستوى ثقة مساو لـ 95%، قم بتقدير الوسط الحسابي للمجتمع $\mu$

**حل المثال. 3:**

لدينا مستوى الثقة: $1-\alpha = 95\%$

مستوى الخطأ او المعنوية : 5%

**ومنه: $Z_{1-\alpha/2} = 1.96$**

كما يتضح لنا أن تباين المجتمع $\sigma^2$ مجهول، ولذلك سنستعين بقيمة الانحراف المعياري لتوزيع المعاينة للمتوسطات $\sigma_{\bar{x}i}$ بدل الانحراف المعياري للمجتمع، وحجم العينة مساو لـ 10 أي اصغر من 30 ففي هذه الحالة سنستخدم توزيع ستيودنت لتحديد مجالات الثقة لـ $\mu$.

وبالتالي لتقدير الوسط الحسابي للمجتمع سنستعين بالعلاقة التالية:

$$\bar{X} - t_{(1-\alpha\,2\,,\,n-1)} * \sigma_x / \sqrt{n} < \mu < \bar{X} + t_{(1-\alpha\,2\,,\,n-1)} * \sigma_x / \sqrt{n}$$

$$72 - t_{(1-0.05\,2\,,\,10-1)} * 6.4/\sqrt{10} < \mu < 72 + t_{(1-0.05\,2\,,\,10-1)} * 6.4/\sqrt{10}$$

$$67.42 < \mu < 76.54$$

حيث تم حساب قيمة $t_{(1-\alpha\,2\,,\,n-1)}$ المساوية للقيمة $t_{(1-0.05\,2\,,\,10-1)}$ وذلك حسب معطيات المثال والتي وجد أنها تقابل القيمة : 2.262 في الجدول الاحصائي لتوزيع ستودنت ويصبح لدينا مجال الثقة كالتالي:

**[ 76.54؛ 67.42]**

إذا نقول أن متوسط مداخيل أفراد تلك الدولة يتراوح ما بين **76.54** كحد أعلى و67.42 كحد أدنى.

# المحاضرة الخامسة:

## نظرية التقدير-الجزء الثاني

### 3.3. فترة الثقة للنسبة

#### 1.3.3. حالة المجتمع غير محدود أو المعاينة غير نفادية والعينة الممتدة ($n \geq 30$) :

لتكن s إحصائية تمثل نسبة "نجاحات" في عينة ذات حجم $n \geq 30$ مستخرجة من مجتمع ثنائي حيث p هي نسبة النجاحات. تستعمل التوزيع الطبيعي لتقدير p فنعين حدود الثقة لـ p كما يلي: $\dot{p} \pm z_c\sigma_p$ أين $\dot{p}$ نسبة النجاحات في العينة.

نعلم من المحور السابق أن $\sigma_p = \sqrt{\frac{pq}{n}}$ ومنه يحدد مجال الثقة لـ p كما يلي (صالح، 2006، 94):

$$\mathbf{ICp=}\ p' \pm z_c \cdot \sqrt{\frac{p(1-p)}{n}}$$

#### 2.3.3.حالة كون المجتمع محدود ذا حجم N والمعاينة نفادية:

$$\mathbf{ICp=}\ p' \pm z_c \cdot \sqrt{\frac{p(1-p)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

**مثال.01:**

أخذت عينة عشوائية من احدى المجتمعات تتكون من ذكور واناث حجمها 85 وجد منها 10 ذكور.

**المطلوب:**

أوجد فترة الثقة لنسبة الذكور في المجتمع P عند فترة الثقة 95% .

**الحل:**

لدينا:

مستوى الثقة: $1-\alpha = 95\%$

مستوى الخطأ او المعنوية : 5%

**ومنه: $Z_{1-\alpha/2} = 1.96$**

n= 85

$\dot{p}=10/85=0.12$

بالتعويض في مجال الثقة التالي نجد:

$$\mathbf{ICp=}\ p' \pm z_c \cdot \sqrt{\frac{p(1-p)}{n}}$$

$$\mathbf{ICp=}0.12 - 1.96*\sqrt{0.12(1-0.12)/85} \leq p \leq 0.12 + 1.96*\sqrt{0.12(1-0.12)/85}$$

$$\mathbf{ICp=}0.05 \leq p \leq 0.19$$

أي أنه عند مستوى ثقة مقدارها 95% ستكون نسبة الذكور في المجتمع محل الدراسة تقع ضمن المجال [0.05 , 0.19]

### 4.3. فترة الثقة للفرق بين وسطين

#### 1.4.3. حالة تباييني المجتمعين معلوم

اذا كان لدينا مجتمعين يتبعان التوزيع الطبيعي وفق العلاقة التالية (السقاف، 2020، ص.129):

$$\bar{X}_1 - \bar{X}_2 \approx N(\mu_{\bar{x}1\,-\,\bar{x}2}\ ;\ \sigma^2_{\bar{X}1\,-\,\bar{X}2})$$

فتكون فترة الثقة (1-α) % كالتالي:

$$-Z_{1-\frac{\alpha}{2}} < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < Z_{1-\frac{\alpha}{2}}$$

و بعد تبسيط المتباينة نجد أن:

$$(\bar{X}_1 - \bar{X}_2) - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < (\mu_1 - \mu_2) < (\bar{X}_1 - \bar{X}_2) + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

ويصبح مجال الثقة كالتالي:

$$IC\mu_1 - \mu_2 = \left[(\bar{X}_1 - \bar{X}_2) \mp Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

### 2.4.3. حالة تباييني المجتمعين مجهول مع حجمي العينة ( n1,n2 ≥30)

اذا كان تبايني مجتمعين $\sigma_2^2$ ؛ $\sigma_1^2$ مجهولين وكان حجم العينتين المأخوذتين من المجتمعين ( n1,n2 ≥30)عندئذ يبقى التوزيع المستخدم هو التوزيع الطبيعي المعياري مع تبديل تبايني المجتمعين المجهولين بتبايني العينة وبالتالي تصبح العلاقة كما يلي (صالح، 2006، ص.97): $\sigma_2^2$؛ $\sigma_1^2$

$$(\bar{X}1 - \bar{X}2) - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_{\bar{x}1}^2}{n_1} + \frac{\sigma_{\bar{x}2}^2}{n_2}} < (\mu_1 - \mu_2) < (\bar{X}1 - \bar{X}2) + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_{\bar{x}1}^2}{n_1} + \frac{\sigma_{\bar{x}2}^2}{n_2}}$$

ويصبح مجال الثقة كالتالي:

$$IC\mu_1 - \mu_2 = \left[(\bar{X}1 - \bar{X}2) \mp Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_{\bar{x}1}^2}{n_1} + \frac{\sigma_{\bar{x}2}^2}{n_2}}\right]$$

### 3.4.3. حالة تباييني المجتمعين مجهول مع حجمي العينة ( n1,n2 <30) أو أحدهما

عندما يكون لدينا مجتمعين احصائيين غير معلومي التباين ولكن توزيعها التكراري من النوع الطبيعي ويكون حجم العينتين المأخوذتين من هذين المجتمعين صغيرا ( n1,n2 <30) ، فإننا نستخدم في هذه الحالة متباينة الثقة الخاصة بتوزيع t لـ ستيودنت حيث يتم حساب درجة الحرية وفق هذه للتوزيع كالتالي (السقاف، 2020، ص.130):

$$V=n_1+n_2-2$$

ويقع ضمن متباينة الثقة المدونة أدناه وذلك بدرجة ثقة: P= (1- α)%

ويصبح مجال الثقة كالتالي:

$$IC\mu_1 - \mu_2 = \left[(\bar{X}_1 - \bar{X}_2) \mp t_{\left(1-\frac{\alpha}{2},\ n_1+n_2-2\right)}\sqrt{\frac{\sigma_{\bar{x}1}^2}{n_1} + \frac{\sigma_{\bar{x}2}^2}{n_2}}\right]$$

**مثال.01.**

قمنا بسحب عينة عشوائية حجمها $n_1$ = 80 ومتوسطها $\bar{X_1}$=1680 وعينة عشوائية ثانية حجمها n2 =75 ومتوسطها $\bar{X_2}$=1200 وكان الانحراف المعياري للمجتمعين على التوالي؛ $\sigma_1^2$=500؛ $\sigma_2^2$=600.

**المطلوب:**

بمستوى ثقة مساو لـ 99%، قم بتقدير فترة الثقة للفرق بين الوسط الحسابي للمجتمعين $\mu_1$-$\mu_2$

**حل المثال. 01:**

لدينا مستوى الثقة: $1-\alpha = 99\%$

مستوى الخطأ او المعنوية : $1\%$

**ومنه:** $Z_{1-\alpha/2} = 2.58$

ومنه فإن فترة الثقة للفرق بين متوسطي المجتمعين بعد التعويض يعطى بالعلاقة التالية:

$$(1680 - 1200) - Z_{1-\frac{0.01}{2}}\sqrt{\frac{500^2}{80}+\frac{600^2}{75}} < (\mu 1 - \mu 2) < (1680 - 1200) + Z_{1-\frac{0.01}{2}}\sqrt{\frac{500^2}{80}+\frac{600^2}{75}}$$

$$250 < \mu 1 - \mu 2 < 710$$

أي عند مستوى ثقة 99% ستكون قيمة الفرق بين متوسطي المجتمعين ضمن المجال:

$$[710 - 250]$$

**مثال.02.**

اذا قمنا بسحب عينة عشوائية حجمها $n_1 = 8$ ومتوسطها $\overline{X_1}=1680$ وعينة عشوائية ثانية حجمها $n2 = 12$ ومتوسطها $\overline{X_2}=1200$ وكان الانحراف المعياري لمتوسطي توزيع المعاينتين على التوالي كما يلي؛ $\sigma_{\overline{x}1}{}^2=500$؛ $\sigma_{\overline{x}2}{}^2=600$.

**المطلوب:**

بمستوى ثقة مساو لـ 99%، قم بتقدير فترة الثقة للفرق بين الوسط الحسابي للمجتمعين $\mu_1$-$\mu_2$

**حل المثال. 02:**

بتطبيق العلاقة التالية:

$$IC\mu_1 - \mu_2 = \left[(\overline{X}_1 - \overline{X}_2) \mp t_{\left(1-\frac{\alpha}{2},\ n_1+n_2-2\right)}\sqrt{\frac{\sigma^2_{\overline{x}1}}{n_1}+\frac{\sigma^2_{\overline{x}2}}{n_2}}\right]$$

$$t_{\left(1-\frac{0.01}{2},\ 8+12-2\right)} = t_{(0.995,\ 18)} = 2.878$$

مع:

ومنه فإن فترة الثقة للفرق بين متوسطي المجتمعين بعد التعويض يعطى بالعلاقة التالية:

$$223.79 < \mu 1 - \mu 2 < 736.20$$

أي أنه عند مستوى ثقة 99% ستكون قيمة الفرق بين متوسطي المجتمعين ضمن المجال:

**[736.20 – 223.79]**

**5.3. فترة الثقة للتباين:**

لتقدير التباين والانحراف المعياري لمجتمع بمجال ثقة نستعمل الخاصية (صالح، 2006، ص.94): $\frac{nS^2}{\sigma^2} = \frac{(n-1)\hat{S}^2}{\sigma^2} \sim \chi^2_{n-1}$.

**مثال.01:** مستوى الثقة يحدد بـ 95 % كما يلي:

$$\chi^2_{0.025} \le \frac{nS^2}{\sigma^2} = \frac{(n-1)\hat{S}^2}{\sigma^2} \le \chi^2_{0.975}$$

ومنه نستنتج مجال الثقة لـ σ كما يلي:

$$\frac{\sqrt{n}S}{\chi_{0.975}} \le \sigma \le \frac{\sqrt{n}S}{\chi_{0.025}} \qquad ou \qquad \frac{\sqrt{n-1}\hat{S}}{\chi_{0.975}} \le \sigma \le \frac{\sqrt{n-1}\hat{S}}{\chi_{0.025}}$$

نظرا لأن توزيع ك2 غير متماثل فإن المجال أعلاه ليس الأمثل، إذ توجد طريقة لتضييق مجال الثقة أكثر إذا لم نشأ أن تكون أطراف المنحنى متساوية، وهذا بخلاف التوزيعات المتماثلة كالطبيعي وستيودنت.

**مثال.02:**

أوجد فترة الثقة 95 % لتباين مجتمع سحبت منه عينة حجمها 8 وتباينها 12.6

**الحل:**

لدينا من المعطيات ما يلي:

$$n=8 \; ; \quad S^2=12.6 \; ; \quad \alpha=5\%$$

من جدول توزيع كاي تربيع نجد:

$$X^2_{(1-\frac{0.05}{2},\ 8-1)} = X^2_{(0.975,\ 7)} = 16$$

$$X^2_{(\frac{0.05}{2},\ 8-1)} = X^2_{(0.025,\ 7)} = 1.69$$

بالتعويض في العلاقة نجد:

$$(8-1)*12.6/16 < \sigma^2 < (8-1)*12.6/1.69$$

$$5.51 < \sigma^2 < 52.18$$

**6.3. فترة الثقة لنسبة تباينين** (صالح، 2006، ص.95):

رأينا مما سابقا أنه إذا كان لدينا مجتمعان طبيعيان تبايناهما $\sigma^2_1$ , $\sigma^2_2$ وسحبنا منهما عينتين عشوائيتين حجمهما على التوالي $n_1$ , $n_2$ فإن : $F = \dfrac{\left[\dfrac{S_1^2 n_1}{n_1-1}\right]\dfrac{1}{\sigma_1^2}}{\left[\dfrac{S_2^2 n_2}{n_2-1}\right]\dfrac{1}{\sigma_2^2}} = \dfrac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2} \rightarrow F_{n_1-1;\ n_2-1}$

إذا يمكن تكوين تقدير بمجال لـ F عند مستوى ثقة 0.98 كما يلي:

$$F_{0.01} \le \frac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2} \le F_{0.99}$$

و من ثم يمكن تقدير النسبة بين تباييني المجتمعين كما يلي:

$$\frac{1}{F_{0.99}}\frac{\hat{S}_1^2}{\hat{S}_2^2} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{1}{F_{0.01}}\frac{\hat{S}_1^2}{\hat{S}_2^2}$$

**مثال.01:**

سحبت عينة عشوائية حجمها n1=8 وتباينها S2=82 من مجتمع يتبع التوزيع الطبيعي، و سحبت عينة عشوائية أخرى مستقلة عن الأولى حجمها n2=11 وتباينها S2=160 من مجتمع يتبع التوزيع الطبيعي.

**المطلوب:**

ماهي حدود فترة الثقة لتباييني المجتمعين باعتماد مستوى ثقة مساو لـ 90%

**الحل:**

بالتعويض في علاقة فترة الثقة لنسبتي تبايني مجتمعين نجد:

$$\frac{82}{160}.F_{(8-1,11-1)}^{\frac{0.1}{2}} < \frac{{\sigma_1}^2}{{\sigma_2}^2} < \frac{82}{160}.F_{(8-1,11-1)}^{1-\frac{0.1}{2}}$$

$$2.49\ .0.28 < \frac{{\sigma_1}^2}{{\sigma_2}^2} < 2.49\ .\ 3.14$$

حيث:

$$F_{(7,10)}^{0.95} = 3.14$$

$$F_{(7,10)}^{0.05} = \frac{1}{F_{(10,\ 7)}^{0.95}} = \frac{1}{3.57} = 0.28$$

وهو ما ينتج عليه ما يلي:

$$0.6972 < \frac{{\sigma_1}^2}{{\sigma_2}^2} < 7.81$$

أي أنه عند مستوى ثقة 95 % ستقع قيمة نسبة التباين لمجتمعين ضمن المجال [ 7.81 ؛ 0.6972]

يمكننا تلخيص الحالات السابقة لبناء فترة الثقة من خلال الجدول الثالثي:

**الجدول.06: فترة الثقة حسب معلومية التباين والتوزيع الاحصائي للنسبة، للتباين وللنسبة بين تباينين**

| المجتمع | التوزيع الاحتمالي للإحصائية | مجال الثقة |
|---|---|---|
| مجتمع غير محدود أو معاينة غير نفادية و عينة ممتدة ($n \geq 30$) | التوزيع الطبيعي | $p' \pm z_c \cdot \sqrt{\frac{p(1-p)}{n}}$ |
| مجتمع محدود ذا حجم **N** والمعاينة نفادية | التوزيع الطبيعي | $p' \pm z_c \cdot \sqrt{\frac{p(1-p)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$ |
| غير معلوم | $\frac{nS^2}{\sigma^2} = \frac{(n-1)\hat{S}^2}{\sigma^2} \sim \chi^2_{n-1}$ | $\frac{\sqrt{n}S}{\chi_{0.975}} \leq \sigma \leq \frac{\sqrt{n}S}{\chi_{0.025}}$ أو $\frac{\sqrt{n-1}\hat{S}}{\chi_{0.975}} \leq \sigma \leq \frac{\sqrt{n-1}\hat{S}}{\chi_{0.025}}$ |
| مجتمعين طبيعيين، أو عينتين مسحوبتين من مجتمع طبيعي واحد. | $F = \frac{\left[\frac{S_1^2 n_1}{n_1-1}\right]\frac{1}{\sigma_1^2}}{\left[\frac{S_2^2 n_2}{n_2-1}\right]\frac{1}{\sigma_2^2}} = \frac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2} \rightarrow F_{n_1-1;\ n_2-1}$ | مثلا عند مستوى ثقة 0.98: $\chi^2_{0.025} \leq \frac{nS^2}{\sigma^2} = \frac{(n-1)\hat{S}^2}{\sigma^2} \leq \chi^2_{0.975}$ $\frac{1}{F_{0.99}}\frac{\hat{S}_1^2}{\hat{S}_2^2} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{1}{F_{0.01}}\frac{\hat{S}_1^2}{\hat{S}_2^2}$ |

**المصدر:** من اعداد الباحث اعتمادا على العلاقات الواردة في النظريات السابقة

# المحاضرة السادسة:

## اختبار الفرضيات-الجزء الأول

## 4. اختبار الفرضيات

تناولنا في المحور الثالث من هذه المطبوعة كيفية تقدير معالم المجتمع من خلال بيانات العينة وبعض خصائص المقدر الجيد؛ وبما أن الدارس يحتاج أحيانا في مرحلة ما من بحثه إلى اختبار فرضية أو أكثر بخصوص المجتمع المدروس. فارتئينا في هذا المحور الرابع تناول كيفية اختبار فرضيات موضوعة حول معالم مجتمع أو أكثر باعتبارها تشكل أحد المواضيع الأساسية للاستدلال الإحصائي، وهذا بغية الوصول إلى قرار بشأن معلمة المجتمع من خلال قبول او رفض تقديرها المعتمد على معطيات العينة المسحوبة من ذلك المجتمع، ففي غالب الأحيان يحاول الباحث اتخاذ قرار لمشكلة محددة بشأن خواص توزيع ما (متوسط او نسبة لعينة عشوائية تم سحبها من المجتمع) ، معتمدا على قرار احصائي لذا يتوجب عليه وضع فروض عن خواص المجتمع ، ومن هنا يختبر مدى صحة هذا الفرض؛ ومن أمثلة ذلك: اختبار فرضية بخصوص معدل الدخل في منطقة معينة، اختبار فرضية نسبة شفاء لدواء معين، ... حيث يتم ذلك بصياغة فرضية عن المجتمع المدروس (أو المجتمعات المدروسة) ومن ثم محاولة الحصول على دليل إحصائي ينفي أو يثبت هذه الفرضية وذلك من خلال بيانات عينة (أو أكثر) عشوائية بسيطة. أين تخص الفرضية أحد معالم المجتمع كالمتوسط، النسبة أو التباين، ونعتمد في إثباتها أو رفضها على خصائص إحصائية معاينة مختارة؛ من أجل ذلك يعتمد هذا الدرس، كما هو الحال بالنسبة لدرس التقدير، على درس المعاينة، أين سنتناول في متن هذا المحور مختلف هذه التفاصيل وكذا مراحل اختبارات الفروض بالاعتماد على العناصر التالية:

### 1.4. ضبط بعض المفاهيم الأساسية حول اختبارات الفروض

#### 1.1.4. تعريف اختبارات الفروض

يعتبر اختبار الفروض أحد أساليب الاحصاء الاستدلالي الذي يستخدم فيه بيانات العينة المسحوبة من مجتمع الدراسة لاتخاذ قرارات أو اصدار أحكام حول هذا المجتمع، وعليه فإن اختبار الفرضيات الاحصائية هو اختبار صحة وصدقية القيمة (القيم) التي تأخذها المعلمة (المعلمات) التي تم تقديرها، وعادة ما تجرى الاختبارات حول قيمة محددة، أو حول النطاق (الفترة) التي وقعت ضمنها أو خارجها (موراي، جون، و ألو، 2004، ص.94).

#### 2.1.4. تعريف الفرضية الاحصائية

نعرف بأنها عبارة عن ادعاء أو تخمين معين حول معلمة من معالم المجتمع، وهذا الفرض قابل لأن يكون صحيحا أو غير صحيح، ولا يمكن التأكد من ذلك إلا من خلال اختباره احصائيا (موراي، جون، و ألو، 2004، ص.94).

#### 3.1.4. أنواع الفرضيات

يجب أن تكون الفرضيات متكاملة بحيث تشمل كل النتائج الممكنة وتبادلية بحيث تلغي احداهما الأخرى وتنقسم الى:

**أ. الفرضية الصفرية أو فرضية العدم**

وهي الفرضية الأساسية المراد اختبارها حول معلمة المجتمع التي نجري عليها الاختبار باستخدام بيانات العينة، والتي يرمز لها بالرمز $H_0$، والتي تشير إلى عدم وجود فرق بين معلمة المجتمع واحصاءة العينة المسحوبة منه (السقاف، 2020، ص.137).

**ب. الفرضية البديلة**

في اختبارات الفروض يتحتم وضع فرض آخر غير فرض العدم المراد اختباره والذي يطلق عليه اصطلاحا بالفرض البديل،

الذي سيتم قبوله في حالة رفض فرض العدم، وبالتالي يمكن القول بأنها هي الفرضية البديلة لفرضية العدم، ويتم قبولها عند رفض الفرضية الصفرية $H_0$ والتي يرمز لها بالرمز $H_1$ .

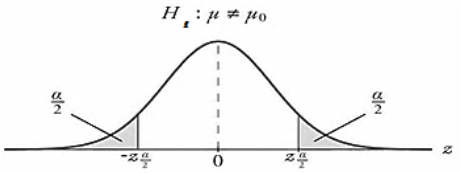
تجدر بنا الاشارة هنا بأن الفرض البديل هو الذي يحدد نوع الاختبار، حيث نميز بين ثلاثة أشكال من الاختبارات، الأول اختبار ثنائي أو اختبار من طرفين والثاني اختبار أحادي جهة اليمين والثالث اختبار أحادي جهة اليسار (السقاف، 2020، ص.138).

**ج. الاختبار الثنائي:**

في هذه الحالة الفرض البديل له طرفان واحد على اليمين والآخر على اليسار حيث نقبل فرضية العدم إذا وقعت قيمة دالة الاختبار في منطقة قبول $H_0$ ونرفض$H_0$ إذا وقعت قيمة دالة الاختبار في منطقة رفض $H_0$ وهي المنطقة المظللة، ويكتب كمايلي (موراي، جون، و ألو، 2004، ص.101):

$$H_0 : \mu = \mu_0$$
$$H_0 : \mu \neq \mu_0$$

ويمكن تمثيله كما هو موضح في الشكل الموالي:




وتسمى المنطقة المحصورة بين ($z_{\alpha/2}$) و ($-z_{\alpha/2}$) بمنطقة القبول ( $-z_{\alpha/2} < z < z_{\alpha/2}$)؛ بينما تسمى المناطق المظللة ( $z > z_{\alpha/2}$) و ( $z < -z_{\alpha/2}$) بمنطقتي الرفض؛ بينما تسمى $z_{\alpha/2}$ و $-z_{\alpha/2}$ بالقيم الحرجة.

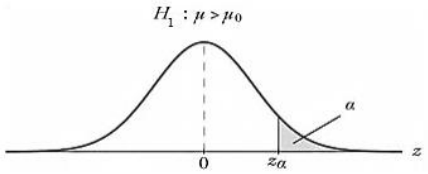
**د. الاختبار أحادي الاتجاه جهة اليمين أو جهة اليسار:**

ويكون في حالة اختبار من جهة واحدة، جهة اليمين أو جهة اليسار (موراي، جون، و ألو، 2004، ص.101):

- **ففي حالة اختبار من جهة اليمين فيصاغ الفرض كما يلي :**

$$H_0 : \mu = \mu_0$$
$$H_0 : \mu > \mu_0$$

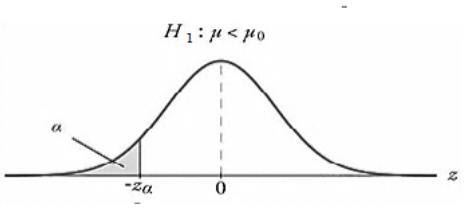
حيث أن منطقة القبول في هذه الحالة هي ( $z < z_\alpha$ )؛ كما هو موضح في الشكل الموالي:




- **ففي حالة اختبار من جهة اليسار فيصاغ الفرض كما يلي :**

$$H_0 : \mu = \mu_0$$
$$H_0 : \mu < \mu_0$$

حيث أن منطقة القبول في هذه الحالة هي ( $z > -z_{\alpha/2}$ )؛ كما هو موضح في الشكل الموالي:




### 4.1.4. أنواع الأخطاء الاحصائية عند اتخاذ القرار:

ان عملية اتخاذ أي قرار إحصائي تنطوي على أخطاء بنسب معينة، حيث أنه من المحتمل أن نرفض فرضية معينة في حين أنها صحيحة، والعكس صحيح، وعموما يمكن القول أنه هناك نوعان من الأخطاء الإحصائية وهي (السقاف،2020، ص.ص، 139.138):

**ا-الخطأ من النوع الاول:**

هو الخطأ الذي نقع فيه عندما نرفض فرضية العدم $H_0$ عندما تكون صحيحة ويرمز لاحتمال وقوع هذا الخطأ بالرمز $\alpha$ونسميه مستوى دلالة الاختبار أو مستوى معنوية الاختبار .

**ب-الخطأ من النوع الثاني:**

هو الخطأ الذي نقع فيه عندما نقبل الفرضية الصفرية بالرغم من عدم صحتها، ويرمز إلى احتمال هذا الخطأ بالرمز $\beta$ ، حيث يمكن تحديد عدة مستويات لذلك ولكن جرت العادة على استخدام واعتماد مستويين وهما 5% و 1% كما هو موضح في الشكل أدناه:

**الجدول.07 : يبين رموز احتمالات الوقوع في الخطأ من النوع الاول $\alpha$ والثاني $\beta$**

| القرار / الحالة | $H_0$ صحيحة | $H_0$ خاطئة |
|---|---|---|
| قبول $H_0$ | قرار سليم | $\beta$ |
| رفض $H_0$ | $\alpha$ | قرار سليم |

**المصدر:** علي أحمد السقاف، **الاحصاء الوصفي والاستدلالي**، المركز الديمقراطي العربي، برلين، ألمانيا، الطبعة الأولى، 2020، ص.139

علما بأن العلاقة بين $\alpha$ و $\beta$ يمكن أن تفسر على النحو التالي:

- ✓ نقصان في أحدهما يزيد الأخرى
- ✓ زيادة حجم العينة n يقلل من احتمال الوقوع في كلا الخطأين

### 2.4. اختبارات الفروض حول الوسط الحسابي للمجتمع

يتناول هذا الاختبار متوسط المجتمع ($\mu$)، مثل متوسط الدخل، متوسط وزن منتج معين، .. حيث يؤكد اختبار الوسط الحسابي فرضية مساواته لقيمة ما ممثلة بـ $\mu_0$؛ للقيام بهذا الاختبار نسحب عينة عشوائية ثم نحسب قيمة متوسطها الحسابي $\overline{X}$ ثم نستخدم التوزيع الاحتمالي لـ $\overline{x}$ من أجل قياس مدى قرب أو بعد هذه القيمة عن $\mu_0$؛ حيث سنتطرق في هذا العنصر الى كل من:

- ✓ الاختبار ثنائي الاتجاه حول الوسط الحسابي للمجتمع
- ✓ الاختبار أحادي الاتجاه حول الوسط الحسابي للمجتمع
- ✓ استخدام S كمقدر لتباين المجتمع في اختبار الوسط الحسابي للمجتمع
- ✓ اختبار الوسط الحسابي للمجتمع باستخدام توزيع t لستيودنت
- ✓

### 1.2.4. اختبار ثنائي الاتجاه حول الوسط الحسابي للمجتمع

سنعتمد المثال الموالي للقيام بعملية اختبار ثنائي الاتجاه حول الوسط الحسابي للمجتمع، والذي من خلاله سنتناول العناصر الأساسية لاجراء هذا الاختبار.

**المثال.01** (صالح، 2006، ص.101)**:**

نريد اختبار فرضية حول الوسط الحسابي لدخل الطالب في السنة الأولى من تخرجه، ولتكن القيمة الافتراضية هي 15000دج كمتوسط للدخل الشهري؛ فللقيام بهذا الاختبار نحتاج إلى الخطوات التالية:

أ.1. تحديد الفرضيات؛

أ.2. تحديد قاعدة القرار؛

أ.3. حساب القيمة الجدولية للمتغيرة؛

أ.4. حساب القيمة الفعلية للمتغيرة؛

أ.5. اتخاذ القرار.

**أ.1. تحديد الفرضيات (الصفرية والبديلة):**

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

تسمى الفرضية $H_0$ **الفرضية الصفرية** أو فرضية العدم، ويؤدي الاختبار إما إلى رفضها ونكتب $RH_0$ وفي هذه الحالة نقبل **الفرضية البديلة** أو عدم رفضها ونكتب $\overline{R}H_0$. $\mu_0$ هي القيمة الافتراضية لـ $\mu$ وهي في هذه الحالة 15000دج كمتوسط للدخل الشهري؛ وبالتالي يمكننا أن نكتب الفرضيات كما يلي:

$$H_0 : \mu = 15000$$
$$H_1 : \mu \neq 15000$$

هنا يجب التنويه على أنه عادة ما تكون قيمة $\mu_0$ محددة بناءا على بيانات عينة عشوائية بسيطة $(\mu_0 = \overline{x})$، وفي هذه الحالة يمكن استخدام الخاصية $\overline{x}_i \sim N(\mu, \sigma^2/n)$ لاجراء هذا الاختبار، حيث أنه تحت $H_0$ فإن : $\overline{x}_i \sim N(\mu_0 , \sigma^2/n)$ مما يعني معلومية احتمال أن يكون $\overline{x}$ قريب إلى درجة ما من $\mu_0$ فمثلا :

$$P(\mu_0 - 1.64(\sigma\overline{x}_i) \leq \overline{x} \leq \mu_0 + 1.64(\sigma\overline{x}_i)) = 0.90$$
$$P(\mu_0 - 1.96(\sigma\overline{x}_i) \leq \overline{x} \leq \mu_0 + 1.96(\sigma\overline{x}_i)) = 0.95$$
$$P(\mu_0 - 2.58(\sigma\overline{x}_i) \leq \overline{x} \leq \mu_0 + 2.58(\sigma\overline{x}_i)) = 0.99$$

وبصفة عامة نكتب:

$$P[\mu_0 - z_{1-\alpha/2}(\sigma\overline{x}_i) \leq \overline{x} \leq \mu_0 + z_{1-\alpha/2}(\sigma\overline{x}_i)] = 1-\alpha$$

أو حسب الكتابة الأكثر شيوعا:

$$P(-z_{1-\alpha/2} \leq \frac{\overline{x} - \mu_0}{\sigma_m} \leq z_{1-\alpha/2}) = 1-\alpha$$

حيث:

- $z \sim N(0, 1)$ حيث، $z_c$ : $(\overline{x} - \mu_0)/ \sigma\overline{x}_i$ (**متغيرة القرار**) هي المتغيرة المعيارية لـ $\overline{x}$ ونرمز لها بـ $z_c$
- $\sigma\overline{x}_i$ تحدد كما يلي:

$\sigma\overline{x}_i = \sigma/\sqrt{n}$ في حالة المعاينة بالإرجاع (أو $n \leq 0.05N$ ) و $\sigma\overline{x}_i = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ في الحالة المعاكسة.

- $1 - \alpha/2$ : المساحة على يسار z .
- n : حجم العينة.

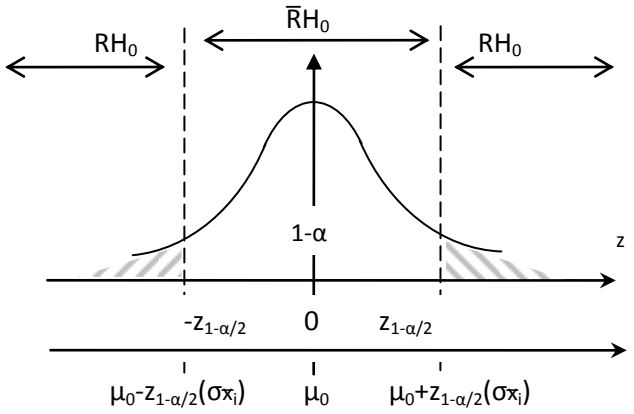
يمكن أن نرفض الفرضية الصفرية التي حدد على أساسها هذا المجال ونقبل بالتالي الفرضية البديلة.إذا كان $\overline{x}$ خارج المجال $1-\alpha$.

تسمى هذه (الخطة) قاعدة القرار.

**أ.2. تحديد قاعدة القرار**

تكتب قاعدة القرار في المثال الذي بين أيدينا، وهي قاعدة **اختبار ثنائي الاتجاه** كما يلي: (أنظر الشكل 1) ،

$$\begin{cases} RH_0 \ si \left| \dfrac{\overline{X} - \mu_0}{\sigma_{\overline{X}}} \right| > z_{1-\alpha/2} \ . \\ \overline{R}H_0 \ \text{sinon.} \end{cases} \quad \text{أو} \quad \begin{cases} RH_0 \ si \ z_c = \dfrac{\overline{X} - \mu_0}{\sigma_{\overline{X}}} \notin \left[ - z_{1-\alpha/2};\ z_{1-\alpha/2} \right] . \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$




**الشكل:** منطقتي القبول والرفض في حالة قاعد القرار الثنائية

تتضمن هذه الخطة مخاطرة تتمثل في الوصول إلى قرار خاطىء:

فقد تكون الفرضية $H_0$ صحيحة بينما تقودنا قيمة $\overline{x}$ المحصلة إلى رفضها؛ وهو ما يسمى هذا **الخطأ من النوع الأول**، كما سبق وأن أشرنا اليه فيما سبق واحتماله $\alpha$، ونكتب :

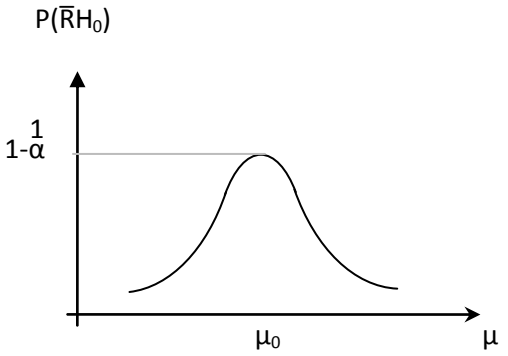
$$P(RH_0 / H_0) = \alpha$$

وقد تقودنا قيمة $\overline{x}$ إلى قبول $H_0$ فيما هي خاطئة، ويسمى هذا **الخطأ من النوع الثاني** واحتماله $\beta$ وتساوي $(1- \alpha)$ ونكتب :

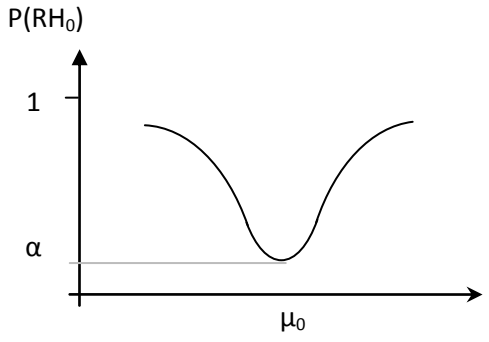
$$P(\overline{R}H_0 / H_1) = 1-\alpha$$

حيث يمكن تقليص احتمال أحد الخطأين على حساب الثاني، ولكن لا يمكن تقليص احتمال كلا الخطأين معا إلا بزيادة حجم العينة.

ويقيس احتمال رفض الفرضية الصفرية $P(RH_0)$ قوة الاختبار (أنظر الشكل 2) فيما يقيس احتمال قبولها $P(\overline{R}H_0)$ فعالية الاختبار (أنظر الشكل 2). ويتوقف كلا الاحتمالين على القيمة الحقيقية لـ $\mu$ .




الشكل (...) منحنى الفعالية




الشكل منحنى القوة (2)

**أ.3. حساب القيمة الجدولية للمتغيرة؛**

يرمز لها بـ $z_t$ حيث، وهي المشار إليها في قاعدة القرار (الشكل الثاني)، وفي حالتنا (اختبار ثنائي بمستوى معنوية 5%) :

$$z_t = z_{1-\alpha/2} = z_{1-0.05/2} = z_{1-0.025} = z_{0.975}$$

ومن جدول التوزيع الطبيعي القياسي نجد أن قيمتها مساوية لـ: $z_{0.975} = 1.96$ .

**أ.4. حساب القيمة الفعلية للمتغيرة؛**

يرمز لها ب $z_c$ وهي المتغير المعيارية لـ $\overline{x}$ (أنظر قاعدة القرار الشكل الأول) :

$$z_c = \frac{\overline{X} - \mu_0}{\sigma_{\overline{X}}} = \frac{15800 - 1500}{1500/\sqrt{100}} = 5.33$$

أ.5. اتخاذ القرار.

بمقارنة القيمة الجدولية والقيمة الفعلية للمتغيرة نقرر قبول أو رفض $H_0$ وذلك حسب قاعدة القرار؛

ففي حالتنا ومثالنا هذا سنرفض $H_0$ لأن القيمة الفعلية أكبر من القيمة الجدولية أي $z_c > z_t$ وبالتالي سنقبل $H_1$ أي أن متوسط دخل الخريج حديث التوظيف لا يساوي 15000دج.

### 2.2.4. الاختبار أحادي الاتجاه حول الوسط الحسابي للمجتمع

يتميز الاختبار الثنائي عن الأحادي في الفرضية البديلة التي هي عدم مساواة في الاختبار الثنائي وأكبر تماما أو أصغر تماما (حسب الحالة) في الاختبار الأحادي، وهذا يترتب عليه تغيير في قاعدة القرار.

**ب.1. الاختبار أحادي الاتجاه من اليمين.**

بالرجوع إلى مثالنا السابق مع تغيير محدد هو أننا نريد اختبار ما إذا كان متوسط دخل الخريج مساو لـ 15000دج أو أكثر (اختبار من جهة اليمين) (السقاف، 2020، ص.140).

ب.1.1 **الفرضيات :**

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0$$

في هذه الحالة $\mu_0 = 15000$ لذلك نكتب :

$$H_0 : \mu = 15000$$
$$H_1 : \mu > 15000$$

ب2.1. **قاعدة القرار**:

$$\begin{cases} RH_0 \ si \dfrac{\overline{X}-\mu_0}{\sigma_{\overline{X}}} > z_{1-\alpha} \\ \overline{R}H_0 \ \text{sinon.} \end{cases} .$$

ب.3.1. **حساب z الجدولية:** (اختبار على اليمين بمستوى معنوية 5%) :

$$z_t = z_{1-\alpha} = z_{1-0.05} = z_{0.95}$$

ومن جدول التوزيع الطبيعي القياس نجد أن: $z_{0.95} = 1.645$

ب.4.1. **حساب z الفعلية**:

$$z_c = \frac{\overline{X}-\mu_0}{\sigma_{\overline{X}}} = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} = \frac{15800-1500}{1500/\sqrt{100}} = 5.33$$

ب.5.1. **القرار**: نرفض $H_0$ لأن $z_c > z_t$ ونقبل $H_1$ أي أن متوسط دخل الخريج حديث التوظيف ليس 15000دج إنما هو أكبر من ذلك.

**ب.2. الاختبار أحادي الاتجاه من اليسار** (السقاف، 2020، ص.141).

بالرجوع إلى مثالنا السابق مع افتراض أن متوسط العينة كان 14200دج ونريد أن نختبر فيما إذا كان متوسط الدخل مساوي لـ 15000دج أم أقل من ذلك

ب.1.2. **الفرضيات :**

$$H_0 : \mu = 15000$$
$$H_1 : \mu < 15000$$

ب.2.2. **قاعدة القرار:**

$$\begin{cases} RH_0 \ si \dfrac{\overline{X}-\mu_0}{\sigma_{\overline{X}}} < -z_{1-\alpha} \\ \overline{R}H_0 \ \text{sinon.} \end{cases} .$$

ب.3.2. **حساب z الجدولية**: (اختبار على اليسار بمستوى معنوية 5 %) :

$$z_t = -\ z_{1-\alpha} = -\ z_{1-0.05} = -\ z_{0.95} = -1.645$$

ب.4.2. **حساب z الفعلية:**

$$z_c = \frac{\overline{X}-\mu_0}{\sigma_{\overline{X}}} = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} = \frac{14200-1500}{1500/\sqrt{100}} = -5.33$$

ب.5.2. **القرار**:

نرفض $H_0$ لأن $z_c < z_t$ ونقبل $H_1$ أي أن متوسط دخل الخريج حديث التوظيف أقل من 15000دج .

**3.2.4. استخدام S كمقدر لـ σ في اختبار الوسط الحسابي للمجتمع** (السقاف، 2020، ص.142):

في مثالنا السابق افترضنا أن قيمة الانحراف المعيار للمجتمع σ معلوم، ولكن في الواقع غالبا ما يكون الانحراف المعياري للمجتمع مجهولا ونحتاج بالتالي إلى استخدام احصاءة الانحراف المعياري للعينة (S) عند حساب $\sigma_{\overline{X}_i}$ (أنظر درس التقدير)، حيث نعوض العبارة $\sigma_{\overline{X}_i} = \sigma/\sqrt{n}$

بـ: $\sigma_X = \dfrac{S}{\sqrt{n-1}}$ أو $\sigma_X = \dfrac{S'}{\sqrt{n}}$

بالمحافظة دائما على مثالنا السابق نفترض أن الانحراف المعياري للدخل الشهري للطالب مجهول، لكن الانحراف المعياري للعينة المسحوبة مساو لـ S = 1600.

**المطلوب:** كيف يمكن اختبار ما إذا كان الدخل الشهري أقل من 15000دج؟

للقيام بهذا الاختبار سنعتمد نفس الخطوات السابقة أثناء قيمنا بالاختبار أحادي الاتجاه من اليسار فيما عدا عملية حساب zالفعلية التي سيكون هناك تغيير في طريقة حسابها، وكذا عملية اتخاذ القرار وذلك كما يلي:

أ.4. **حساب z الفعلية**:

$$z_c = \frac{\bar{X} - \mu_0}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu_0}{S/\sqrt{n-1}} = \frac{14200 - 1500}{1600/\sqrt{99}} = -4.97$$

أ.5. **القرار**:

نرفض $H_0$ لأن $z_c < z_t$ ونقبل $H_1$ أي أن متوسط دخل الخريج حديث التوظيف ليس 15000دج وإنما هو أقل من ذلك.

### 4.2.4. اختبار الوسط الحسابي للمجتمع باستخدام توزيع t لستيودنت (صالح، 2006، ص.104):

في حالة كانت مفردات العينة أقل من 30 أي أن $n< 30$ و كان $\sigma$ (الانحراف المعياري للمجتمع) مجهولا، هنا لا يمكننا استخدام التوزيع الطبيعي، وانما سنعتمد على توزيع t لستيودنت ويصبح بذلك لدينا :

$$\frac{\bar{X} - \mu}{S/\sqrt{n-1}} \sim t_{n-1}$$

وتحت $H_0$ ( $\mu = \mu_0$) :

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n-1}} \sim t_{n-1}$$

يمكن إذا استخدام توزيع ستيودنت (بشرط أن يكون توزيع المجتمع طبيعيا أو على الأقل جرسي الشكل) .

وتتغير قاعدة القرار تبعا لهذا التغيير فتكتب في حالة الاختبار الثنائي كما يلي:

$$\begin{cases} RH_0 \ si \left| \dfrac{\bar{X} - \mu_0}{S/\sqrt{n-1}} \right| > t_{n-1;\, 1-\alpha/2} \\ \bar{R}H_0 \ \text{sinon.} \end{cases} .$$

في حالة اختبار من اليمين:

$$\begin{cases} RH_0 \ si \dfrac{\bar{X} - \mu_0}{S/\sqrt{n-1}} > t_{n-1,\, 1-\alpha} \\ \bar{R}H_0 \ \text{sinon.} \end{cases}$$

في حالة اختبار من اليسار:

$$\begin{cases} RH_0 \ si \dfrac{\bar{X} - \mu_0}{S/\sqrt{n-1}} < -t_{n-1;\, 1-\alpha} \\ \bar{R}H_0 \ \text{sinon.} \end{cases}$$

**مثال:**

ادعى أحد المسؤولين في بنك خاص أن معدل فترة الانتظار في كل شباك لا تزيد عن 15 دقيقة، وبهدف اختبار هذا الادعاء تم دراسة عينة مكونة من 10 شبابيك لخدمة الزبائن على مستوى هذه المؤسسة البنكية فوجد أن معدل الانتظار هو 5.16 دقيقة بانحراف معياري مساو لـ 4 دقائق، على فرض أن مدة الانتظار لكل الشبابيك في البنك تتبع التوزيع الطبيعي.

**المطلوب:** اختبر ادعاء البنك عند مستوى معنوية 5%

**الحل:**

أ.1. **الفرضيات :**

$$H_0 : \mu = 15$$
$$H_1 : \mu > 15$$

أ.2. **قاعدة القرار:**

بما أن العينة مسحوبة من مجتمع طبيعي تباينه مجهول، وحجم العينة صغير n=10 وبالتالي لاجراء هذا الاختبار سنعتمد على توزيع t لستيودنت باعتماد قاعدة القرار التالية:

$$\begin{cases} RH_0 \ si \ \dfrac{\overline{X} - \mu_0}{S/\sqrt{n-1}} > t_{n-1,\,1-\alpha} \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

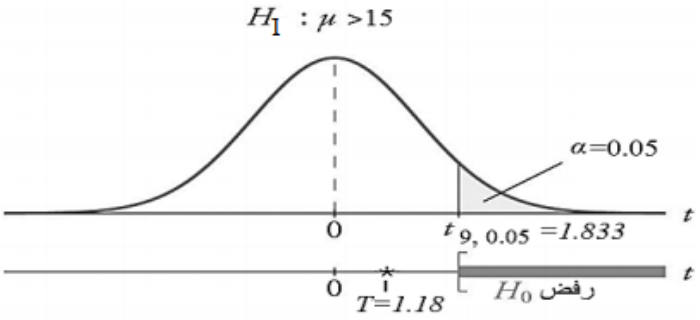
أ.3. **حساب t الجدولية**: (اختبار على اليمين بمستوى معنوية 5 %) :

$$\alpha = 0.05 \Rightarrow t_{n-1;\,\alpha} = t_{9;0.05} = 1.833$$

أ.4. **حساب t الفعلية:**

$$\frac{\overline{X} - \mu_0}{S/\sqrt{n-1}} \sim t_{n-1} \quad \Rightarrow \quad \mathbf{T} = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} = \frac{16.5 \quad -15}{4 \quad /\sqrt{10}} = \ 1.18$$

يمكننا تمثيل منطقة القبول والرفض بيانيا من خلال الشكل الموالي :




أ.5. **القرار**:

بما أن قيمة $t_c$ المحسوبة وقعت في منطقة القبول أي أن $t_c < t_t$ أي أن القرار المتخذ هو قبول فرضية العدم، أي أن ادعاء مسؤول البنك بأن متوسط فترة انتظار العملاء على الشبابيك لا يتعدى 15 دقيقة صحيح عند مستوى معنوية 5%.

### 3.4. اختبارات الفروض حول النسبة

يتعلق هذا الاختبار بنسبة مفردات المجتمع التي تتصف بخاصية ما (p)، حيث يؤكد الاختبار أو ينفي صحة فرضية معينة بخصوص قيمة p يرمز للقيمة الافتراضية بـ $p_0$ وتكتب الفرضية كما يلي (السقاف، 2020، ص.143):

$$H_0 : p = p_0$$

للقيام بهذا الاختبار نستخدم خصائص p' النسبة في العينة (أنظر توزيع المعاينة للنسبة : نظرية 6).

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

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

استنادا إلى هذه الخصائص وتحت H0 : $p' \approx N\left(p_0 ; \sqrt{\frac{p_0 q_0}{n}}\right)$

و من ثم يمكن تحديد قاعدة القرار بحسب طبيعة الاختبار كما يلي:

في حالة الاختبار الثنائي:

$$\begin{cases} RH_0 \ si \left| \frac{p' - p_0}{\sqrt{p_0 q_0 / n}} \right| > z_{1-\alpha/2} \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

في حالة اختبار من اليمين:

$$\begin{cases} RH_0 \ si \ \frac{p' - p_0}{\sqrt{p_0 q_0 / n}} > z_{1-\alpha} \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

في حالة اختبار من اليسار:

$$\begin{cases} RH_0 \ si \ \frac{p' - p_0}{\sqrt{p_0 q_0 / n}} < -z_{1-\alpha} \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

**مثال:**

تقدر الدوائر الرسمية نسبة المتخرجين الجامعيين الذين يحصلون على عمل في السنة الأولى التي تلي تخرجهم بـ 70% ، وجدت دراسة أجريت على عينة من 900 طالب أن نسبة الحصول على عمل في السنة الأولى التي تلت تخرج هؤلاء الطلبة هي 67%.

المطلوب: قم باختبار ما إذا كانت النسبة الرسمية صحيحة أم مبالغ فيها، وذلك عند مستوى معنوية 5%.

للقيام بهذا الاختبار سنتبع الخطوات التالية:

أ.1. **الفرضيات :**

$H_0 : p = 0.70$

$H_1 : p < 0.70$

أ.2. **قاعدة القرار:**

$$\begin{cases} RH_0 \ si \ \frac{p' - p_0}{\sqrt{p_0 q_0 / n}} < -z_{1-\alpha} \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

أ.3. **حساب $z_t$ الجدولية**: (اختبار على اليسار بمستوى معنوية 5 %) :

$\mathbf{Z_t} = -Z_{1-0.05} = -1.64$

أ.4. **حساب $Z_c$ الفعلية:**

$$\mathbf{Z_c} = \frac{p'-p_0}{\sqrt{p_0 q_0 / n}} = \frac{0.67-0.7}{\sqrt{0.7(0.3)/900}} \cong -196.34 < -z_{1-0.05} = -1.64$$

أ.5. **القرار:**

نرفض الفرضية $H_0$ لأن $z_c < z_t$ ونقبل $H_1$ أي أن نسبة المتخرجين الجامعيين الذين يحصلون على عمل في السنة الأولى التي تلي تخرجهم أقل من - 70% .

# المحاضرة السابعة:

## اختبار الفرضيات-الجزء الثاني

**4.4. اختبارات المقارنة بين متوسطي مجنمعين** (اختبار تساوي متوسطي مجتمعين)

يتناول هذا الاختبار مقارنة بين مجتمعين من خلال الوسط الحسابي أي أن الغرض من الاختبار هو تأكيد أو نفي تساوي متوسطي مجتمعين من خلال عينتين عشوائيتين بسيطتين مستقلتين، أين سنركز في هذا العنصر على متغيرة القرار، إذ من السهل على الطالب استنتاج كيفية إتمام الخطوات الأخرى على ضوء ما سبق؛ حيث نفرق هنا بين حالتين الأولى عندما يكون تباين المجتمعين معلومين والثانية عندما يكون تبيان المجتمعين مجهولين. وذلك فيما يلي:

**1.4.4. اختبارات المقارنة بين متوسطي مجنمعين عندما يكون تباينهما معلوم** (السقاف، 2020، ص.150)

إذا كان $\overline{x}_1$ هو الوسط الحسابي لعينة عشوائية حجمها $n_1$ تم سحبها من مجتمع يتوزع توزيعا طبيعيا وسطه الحسابي $\mu_1$ وتباينه $\sigma_1^2$ معلوم، وكان $\overline{x}_2$ هو الوسط الحسابي لعينة عشوائية حجمها $n_2$ تم سحبها على التوالي من مجتمع يتوزع توزيعا طبيعيا وسطه الحسابي $\mu_2$ وتباينه $\sigma_2^2$ معلوم أيضا.

اذا أردنا اختبار الفرضية الصفرية التالية :

$$H_0 : \mu_1 - \mu_2 = \overline{x}_1 - \overline{x}_2$$

مقابل الفرضية البديلة حسب طبيعة الاختبار ثنائي أم أحادي جهة اليمين أو اليسار كما يلي:

- ففي **حالة الاختبار الثنائي** ($H_1 : \mu_1 - \mu_2 \neq \overline{x}_1 - \overline{x}_2$) فاننا نرفض فرضية العدم $H_0$ عند مستوى دلالة $\alpha$ اذا كانت ( $z > z_{\alpha/2}$ أو $z < -z_{\alpha/2}$).
- وفي **حالة الاختبار أحادي جهة اليمين** ($H_1 : \mu_1 - \mu_2 > \overline{x}_1 - \overline{x}_2$) فاننا نرفض فرضية العدم $H_0$ عند مستوى دلالة $\alpha$ اذا كانت ( $z > z_{1-\alpha}$).
- أما في **حالة الاختبار أحادي جهة اليسار** ($H_1 : \mu_1 - \mu_2 < \overline{x}_1 - \overline{x}_2$) فاننا نرفض فرضية العدم $H_0$ عند مستوى دلالة $\alpha$ اذا كانت ( $z < -z_{1-\alpha}$).

مع العلم أن متغيرة القرار تكتب بالشكل التالي:

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

**ملاحظة :**

**عند القيام باختبارات المقارنة بين متوسطي مجنمعين يكون تباينهما معلوم فاننا نعرف حالتين هما:**

- في حالة كان المجتمعين يتوزعان توزيعا طبيعيا فان متغيرة القرار تكتب كما يلي:

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0.1)$$

- في حالة العينتين المسحوبتين من مجتمعين ما تكون مفرداتها أكبر من أو تساوي 30 ($n_1 , n_2 \geq 30$)، فان متغيرة القرار تكتب كما يلي:

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \approx N(0,1)$$

**مثال :**

اذا كانت لدينا نتائج عينتين عشوائيتين مستقلتين مسحوبتين من منطقتين لمقارنة متوسط عمر الناخب فيهما حيث :

$$\overline{x}_{1=} 40 ; \overline{x}_{2=} 34 ; \sigma_1^2 =84 ; \sigma_2^2 =27 ; n_1 = 120 ; n_2 = 90$$

**المطلوب :**

اختبر الفرضية التي مفادها بأن متوسط عمر الناخب في المنطقة الأولى يساوي متوسط عمر الناخب في المنطقة الثانية بمستوى معنوية 5% مقابل الفرض البديل على أنهما غير متساويين.

**الحل :**

**أ.1. الفرضيات:**

$$H_0 : \mu_1 - \mu_2 = \overline{x}_1 - \overline{x}_2$$
$$H_1 : \mu_1 - \mu_2 \neq \overline{x}_1 - \overline{x}_2$$

**أ.2. الاحصائية المحسوبة $Z_C$ تحسب كما يلي:**

$$Z_C = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

بالتعويض نجد أن قيمة الاحصائية المحسوبة $Z_C$ تساوي القيمة 6 أي $Z_C = 6$

**أ.3. حساب القيمة الجدولية للاحصاءة $z_t$**

$$z_t = z_{1-\alpha/2} = z_{1-0.05/2} = z_{1-0.025} = z_{0.975}$$

ومن جدول التوزيع الطبيعي القياسي نجد أن قيمتها مساوية لـ: $z_{0.975} = 1.96$ .

**أ.4. المقارنة والقرار :**

بما أن قيمة الإحصائية $Z_C = 6$ تقع في منطقة الرفض لأن قيمة $z_c > 1.96$ وهو ما يقودنا الى رفض الفرضية الصفرية وقبول الفرضية البديلة أي أن متوسط عمر الناخب في المنطقة الأولى لا يساوي متوسط عمر الناخب في المنطقة الثانية بمستوى معنوية 5%.

**2.4.4. اختبارات المقارنة بين متوسطي مجتمعين** عندما يكون تباينهما مجهول

عندما يكون لدينا مجتمعين احصائيين غير معلومي التباين (تباينهما مجهول) ولكن توزيعها التكراري من النوع الطبيعي أو عندما يكون حجم العينيتن المأخوذتين من هذين المجتمعين صغير (عندما تكون مفرداتها أقل تماما من 30 $(n_1 , n_2 < 30)$)؛ وأردنا اختبار الفرضية الصفرية التالية :

$$H_0 : \mu_1 - \mu_2 = \overline{x}_1 - \overline{x}_2$$

مقابل الفرضية البديلة حسب طبيعة الاختبار ثنائي أم أحادي جهة اليمين أو اليسار كما يلي:

- ففي **حالة الاختبار الثنائي** ($H_1 : \mu_1 - \mu_2 \neq \overline{x}_1 - \overline{x}_2$) فاننا نرفض فرضية العدم $H_0$ عند مستوى دلالة $\alpha$ اذا كانت ( $T > t_{(\alpha/2,\ n1+n2-2)}$ أو $T < -t_{(\alpha/2,\ n1+n2-2)}$).
- وفي **حالة الاختبار أحادي جهة اليمين** ($H_1 : \mu_1 - \mu_2 > \overline{x}_1 - \overline{x}_2$) فاننا نرفض فرضية العدم $H_0$ عند مستوى دلالة $\alpha$ اذا كانت ( $T > t_{(\alpha,\ n1+n2-2)}$).
- أما في **حالة الاختبار أحادي جهة اليسار** ($H_1 : \mu_1 - \mu_2 < \overline{x}_1 - \overline{x}_2$) فاننا نرفض فرضية العدم $H_0$ عند مستوى دلالة $\alpha$ اذا كانت ( $T < -t_{(\alpha/2,\ n1+n2-2)}$).

مع العلم أن متغيرة القرار تكتب بالشكل التالي:

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma^2_{\overline{x}1}}{n_1} + \frac{\sigma^2_{\overline{x}2}}{n_2}}}$$

**ملاحظة :**

**عند القيام باختبارات المقارنة بين متوسطي مجنمعين يكون تباينهما مجهول فاننا نعرف حالتين هما:**

- في حالة كان المجتمعين يتوزعان توزيعا طبيعيا بتباينين متساويين فان متغيرة القرار تكتب كما يلي (السقاف، 2020، ص.152):

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}$$

- في حالة كان مجتمعين ما مفردتيهما أكبر أو تساوي 30، ($n_1 , n_2 \geq 30$) فان متغيرة القرار تكتب كما يلي:

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}} \approx N(0,1)$$

**مثال**:

نسحب من مجتمعين طبيعيين متساويي التباين عينتين حجم الأولى 18 وحجم الثانية 21. ووجدنا النتائج التالية:

$$\overline{X}_1 = 81,\ \overline{X}_2 = 76,\ S^2_1 = 9,\ S^2_2 = 8.$$

**المطلوب:**

أ.1. القيام باختبار تساوي متوسطي المجتمعين بمستوى معنوية 5 % .

$$H_0 : \mu_1 - \mu_2 = \overline{x}_1 - \overline{x}_2$$

$$H_1 : \mu_1 - \mu_2 \neq \overline{x}_1 - \overline{x}_2$$

**أ.2. الاحصائية المحسوبة $T_C$ تحسب كما يلي:**

$$Tc = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{81 - 76}{\sqrt{\frac{18(9) + 21(8)}{18 + 21 - 2}\left(\frac{1}{18} + \frac{1}{21}\right)}} \cong 5.43$$

اذا بالتعويض في جملة المعادلة السابقة نجد نجد أن قيمة الاحصائية المحسوبة $T_C$ تساوي القيمة 5.43

**أ.3. حساب القيمة الجدولية للاحصاءة $T_t$**

ومن جدول التوزيع الاحصائي لستيودنت نجد أن قيمتها مساوية لـ: $t_{0.975;37} \cong 2.336 < 5.43$

**أ.4. المقارنة والقرار :**

بما أن قيمة الإحصائية $T_C = 5.43$ تقع في منطقة الرفض لأن قيمة $T_t = 2.336$ وهو ما يقودنا الى رفض الفرضية الصفرية وقبول الفرضية البديلة أي أن متوسطي المجتمعين غير متساويين عند مستوى معنوية 5%.

**5.4. اختبارات الفروض حول تباين المجتمع وحول تساوي تباييني مجتمعين**

**1.5.4. اختبارات الفروض حول تباين المجتمع** (صالح، 2006، ص.105):

إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا تباينه $\sigma^2$ وأردنا إجراء اختبارات حول المعلمة $\sigma^2$ فإن التوزيع الإحصائي للاختبار المناسب في هذه الحالة هو لتوزيع الإحصائي الذي يعتمد على أفضل مقدر للمعلمة $\sigma^2$ وهو تباين العينة $\sigma_{\bar{x}}^2$ وهذا التوزيع الاحصائي هو المتغير العشوائي $\chi 2$ حيث :

$$\chi^2 = (n-1)\,\sigma^2_{\bar{x}} \,/\, \sigma^2$$

والتوزيع الاحتمالي لهذا المتغير يسمى توزيع كأي تربيع بدرجة حرية $v = n - 1$

فاذا أردنا اختبار الفرضية الصفرية

$$H_0 : \sigma^2 - \sigma_0^2 = 0$$

مقابل الفرضية البديلة حسب طبيعة الاختبار ثنائي أم أحادي جهة اليمين أو اليسار كما يلي:

- ففي **حالة الاختبار الثنائي** $(H_1 : \sigma^2 - \sigma_0^2 \neq 0)$ فاننا نرفض فرضية العدم $H_0$ عند مستوى دلالة $\alpha$ اذا كانت $(\chi 2_{(\alpha/2,\, n-1)} < \chi 2)$ أو $(\chi 2_{(1-\alpha/2,\, n-1)} > \chi 2)$ ؛ توزيع كاي تربيع غير متناظر
- وفي **حالة الاختبار أحادي جهة اليمين** ففي **حالة الاختبار الثنائي** $(H_1 : \sigma^2 > \sigma_0^2)$ فاننا نرفض فرضية العدم $H_0$ عند مستوى دلالة $\alpha$ اذا كانت $(\chi 2_{(1-\alpha,\, n-1)} < \chi 2)$.
- أما في **حالة الاختبار أحادي جهة اليسار** ففي **حالة الاختبار الثنائي** $(H_1 : \sigma^2 < \sigma_0^2)$ فاننا نرفض فرضية العدم $H_0$ عند مستوى دلالة $\alpha$ اذا كانت $(\chi 2_{(1-\alpha,\, n-1)} > \chi 2)$.

**مثال:**

معمل لانتاج الأنابيب البلاستيكية يسوق انتاجه من الانابيب عندما يكون الانحراف المعياري في سمك جدار الأنبوب بحدود المقدار 0.003 سم، قام مفتش وزارة الصناعة بسحب عينة عشوائية من انتاج يوم ما وهو عبارة عن 20 أنبوب فلوحظ أن الانحراف المعياري في سمك جدار الأنبوب كان 0.004 سم؛ هل سيسمح مفتش الصناعة بتسويق الانابيب المنتج في ذلك اليوم؟

ملاحظة: استخدم مستوى معنوية 5%

**الحل:**

أ.1. للقيام بهذا الاختبار عند مستوى معنوية 5 % سنعتمد الفرضيات التالية:

$$H_0 : \sigma^2 - (0.003)^2 = 0$$
$$H_1 : \sigma^2 - (0.003)^2 \neq 0$$

**أ.2. الاحصائية المحسوبة $T_C$ تحسب كما يلي:**

$$\chi^2 = (n-1)\,\sigma^2_{\bar{x}} \,/\, \sigma^2$$
$$\chi^2 = (20-1)\,(0.004)^2 \,/\, (0.003)^2 = 33.78$$

اذا بالتعويض في جملة المعادلة السابقة نجد أن قيمة الاحصائية المحسوبة $\chi^2$ تساوي القيمة 33.78

**أ.3. حساب القيمة الجدولية للاحصاءة** $\chi 2_{(1-\alpha/2,\ n-1)}$ أو $\chi 2_{(\alpha/2,\ n-1)}$

ومن جدول التوزيع الاحصائي لـ $\chi^2$ نجد أن قيمتهما مساوية لـ:

$$(\chi 2_{(0.025,\ 19)}) = 32.852 \quad , \quad (\chi 2_{(0.975,\ 19)}) = 8.907$$

**أ.4. المقارنة:**

$$\chi^2 = 33.78 > (\chi 2_{(0.025,\ 19)}) = 32.852$$

**أ.5. القرار:**

بما أن قيمة الإحصائية المحسوبة $\chi^2$ تساوي القيمة 33.78 تقع في منطقة الرفض لأن قيمة $(\chi 2_{(0.025,\ 19)}) = 32.852$ وهو ما يقودنا الى رفض الفرضية الصفرية وقبول الفرضية البديلة أي أن المصنع مخالف لمواصفات التصنيع وبالتالي لن يسمح مفتش الصناعة بتسويق المنتج.

**2.5.4.اختبارات الفروض حول تباييني مجتمعين**

ان الغرض من الاختبار هو تأكيد أو نفي تساوي تباينا مجتمعين من خلال عينتين عشوائيتين بسيطتين مستقلتين، حيث تكتب الفرضيات (في حالة الاختبار الثنائي)كما يلي (القبانجي و كرمجي، 2012، ص.ص، 280.282):

$$H_0 : \sigma^2_1 = \sigma^2_2$$

يتم تحويلها على الصورة التالية لأجل سهولة التعامل معها:

$$H_0 : \sigma^2_1/\sigma^2_2 = 1$$

وتأخذ الفرضية البديلة احدى الصور:

$$H_1 : \sigma^2_1/\sigma^2_2 \neq 1$$
$$H_1 : \sigma^2_1/\sigma^2_2 > 1$$

ويكون المعيار المستخدم في هاتين الحالتين كما يلي:

$$F = \frac{\hat{S}_1^2}{\hat{S}_2^2} \sim F_{n_1-1;\, n_2-1}$$

ويكون المعيار المستخدم في هاته الحالتين كما يلي:

$$H_1 : \sigma^2_1/\sigma^2_2 > 1$$

$$F = \frac{\hat{S}_1^2}{\hat{S}_2^2} \sim F_{n_2-1;\, n_1-1}$$

**مثال.1.:**

في دراسة لبيان معنوية الفرق بين تباييني أوزان الأرانب بعمر 6 أشهر في المزرعتين A و B، حيث تم سحب عينة عشوائية من كل مزرعة فكانت النتائج بالكيلو غرام كما يلي:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| المزعة A | 1.5 | 1.7 | 1.3 | 1.4 | 1.1 | 1.4 | 1.2 |
| المزرعة B | 1.7 | 1.6 | 1.3 | 1.4 | 1.4 | 1.5 | |

**المطلوب:**

اختبر تساوي تباييني أوزان الأرانب في المزرعتين من عدمه وذلك باعتماد مستوى معنوية 5%

**الحل:**

من الجدول نجد أن:

$$S^2_1 = 0.039 \text{ و } S^2_2 = 0.022$$

أ.1. للقيام بهذا الاختبار عند مستوى معنوية 5 % سنعتمد الفرضيات التالية:

$$H_0 : \sigma^2_1/\sigma^2_2 = 1$$

$$H_1 : \sigma^2_1/\sigma^2_2 \neq 1$$

**أ.2. الاحصائية المحسوبة $F_C$ تحسب كما يلي:**

$$\mathbf{F} = S^2_1 / S^2_2 = 0.039 / 0.022 = 1.773$$

**أ.3. حساب القيمة الجدولية للاحصاءة F:**

ومن جدول التوزيع الاحصائي لـ F نجد أن قيمتهما مساوية لـ: 6.9777

$$\mathbf{F}_{(n1-1\,;\,n2-1)} = \mathbf{F}_{(6\,;\,5)} = 6.9777$$

**أ.4. المقارنة:**

$$\mathbf{F_t}_{(n1-1\,;\,n2-1)} = \mathbf{F_t}_{(6\,;\,5)} = 6.9777 > \mathbf{F_c} = S^2_1 / S^2_2 = 0.039 / 0.022 = 1.773$$

**أ.5. القرار:**

بما أن قيمة الإحصائية المحسوبة $\mathbf{F_c}$ تساوي القيمة 1.773 تقع في منطقة القبول لأن قيمة $\mathbf{F_t}$ تساوي 6.9777 وهو ما يقودنا الى قبول الفرضية الصفرية ورفض الفرضية البديلة أي أنه لا توجد فروق بين تباييني أوزان الأرانب في المزرعتين A و B.

**مثال .2.**

من مجتمع طبيعي التوزيع سحبت عينة عشوائية بحجم 15 مشاهدة بلغ تباينها 36 ، ومن مجتمع آخر طبيعي التوزيع سحبت عينة عشوائية بحجم 18 مشاهدة بلغ تباينها 25 ،
فهل تباين المجتمع الأول مساو لتباين المجتمع الثاني أم أكبر منه اذا تم الاعتماد على مستوى معنوية مساو لـ 10%.

**الحل:**

أ.1. للقيام بهذا الاختبار عند مستوى معنوية 5 % سنعتمد الفرضيات التالية:

$$H_0 : \sigma^2_1/\sigma^2_2=1$$
$$H_1 : \sigma^2_1/\sigma^2_2>1$$

**أ.2. الاحصائية المحسوبة $F_C$ تحسب كما يلي:**

$$F = S^2_1 / S^2_2 = 36 / 25 = 1.44$$

**أ.3. حساب القيمة الجدولية للاحصاءة F:**

ومن جدول التوزيع الاحصائي لـ F نجد أن قيمتهما مساوية لـ: 1.91

$$F_{(n1-1\ ;\ n2-1)} = F_{(14\ ;\ 17)} = 1.91$$

**أ.4. المقارنة:**

$$F_{t\ (n1-1\ ;\ n2-1)} = F_{t\ (14\ ;\ 17)} = 1.91 > F_c = S^2_1 / S^2_2 = 36 / 25 = 1.44$$

**أ.5. القرار:**

بما أن قيمة الإحصائية المحسوبة $F_c$ تساوي القيمة 1.44 تقع في منطقة القبول لأن قيمة $F_t$ تساوي 1.91 وهو ما يقودنا الى قبول الفرضية الصفرية ورفض الفرضية البديلة أي أنه لا توجد فروق بين تباييني المجتمعين الأول والثاني.