الاحصاء

Prof Mustapha
KHALDI

I. رموز ومصطلحات

N : عدد التكرارات (المجموع الكلي)
Q_1 : الربعي الأول
Q_3 : الربعي الثالث
D_1 : العشري الأول
D_9 : العشري التاسع
Med : الوسيط

II. الربعيان Q_1 و Q_3 و العشريان D_1 و D_9

تحديد Q_1، Q_3، D_1 و D_9		
بالترتيب	طبع كمي متقطع	طبع كمي مستمر
❶ نرتب القيم ترتيبا تصاعديا مع تكراراتها ❷ Q_1 هي القيمة التي رتبتها $\geq \frac{N}{4}$ ❸ Q_3 هي القيمة التي رتبتها $\geq \frac{3N}{4}$ ❹ D_1 هي القيمة التي رتبتها $\geq \frac{N}{10}$ ❺ D_9 هي القيمة التي رتبتها $\geq \frac{9N}{10}$	❶* Q_1 هي أول قيمة التي تكرارها المجمع الصاعد أكبر أو يساوي $\frac{N}{4}$ * أو Q_1 هي أول قيمة التي تواترها المجمع الصاعد أكبر أو يساوي $\frac{1}{4}$ ❷* Q_3 هي أول قيمة التي تكرارها المجمع الصاعد أكبر أو يساوي $\frac{3N}{4}$ * أو Q_3 هي أول قيمة التي تواترها المجمع الصاعد أكبر أو يساوي $\frac{3}{4}$	❶ Q_1 هي فاصلة النقطة من منحنى التواتر المجمع الصاعد التي ترتيبها $\frac{1}{4}$ ❷ Q_3 هي فاصلة النقطة من منحنى التواتر المجمع الصاعد التي ترتيبها $\frac{3}{4}$ ❸ D_1 هي فاصلة النقطة من منحنى التواتر المجمع الصاعد التي ترتيبها $\frac{1}{10}$ ❹ D_9 هي فاصلة النقطة من منحنى التواتر المجمع الصاعد التي ترتيبها $\frac{9}{10}$

III. خواص الربعيات

1) الانحراف الربعي $I = Q_3 - Q_1$

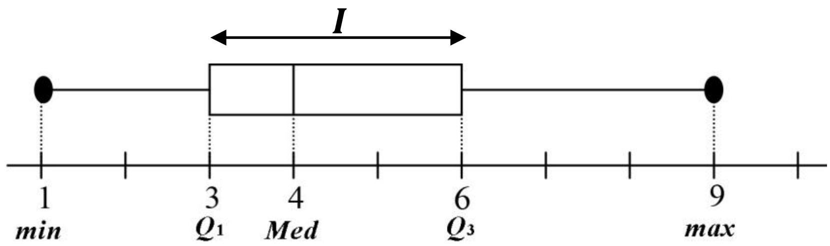
*ملاحظة1: الانحراف الربعي هو مؤشر من مؤشرات التشتت

*ملاحظة2: في مضلع سلسلة مستمرة المستقيمات التي معادلاتها $x = Q_1$ ، $x = Med$ ، $x = Q_3$ تقسم المضلع إلى أربعة مجالات متساوية المساحات

2) المخطط بالعلب

① نضع قيم الطبع على محور (أفقي أو شاقولي)
② نعين على هذا المحور القيم: min ، max ، Q_1 ، Med ، Q_3
③ نكون عندئذ مستطيلا (العلبة) بالتوازي مع المحور بحيث طول المستطيل هو الانحراف الربعي وعرضه كيفي)

مثال: $min = 1$ ، $max = 9$ ، $Q_1 = 3$ ، $Med = 4$ ، $Q_3 = 6$



*ملاحظة: هذا المخطط يمكننا من مشاهدة تشتت توزيع احصائي والمقارنة بين عدة سلاسل احصائية

3) أثر تغيير تآلفي على الربعيين

A سلسلة إحصائية (x_i, n_i) وسيطها Med و ربعيها Q_1 و Q_3

B سلسلة إحصائية (y_i, n_i) بنفس التكرارات وسيطها Med' و ربعيها Q_1' و Q_3'

من أجل كل i لدينا:

- $y_i = ax_i + b$
- $Med' = aMed + b$
- $Q_1' = aQ_1 + b$
- $Q_3' = aQ_3 + b$

حيث $a \in \mathbb{R}^*$ و $b \in \mathbb{R}$

IV. الوسط الحسابي $\bar{X}$

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{p} n_i x_i$$ أو $$\bar{X} = \sum_{i=1}^{p} f_i x_i$$ حيث: $N = \sum_{i=1}^{p} n_i$ و $f_i = \frac{n_i}{N}$

V. التباين V

$$V = \frac{1}{N}\sum_{i=1}^{p} n_i \left(x_i - \bar{X}\right)^2$$ أو $$V = \sum_{i=1}^{p} f_i \left(x_i - \bar{X}\right)$$

أيضا $$V = \frac{1}{N}\sum_{i=1}^{p} n_i x_i^2 - \bar{X}^2$$ أو $$V = \sum_{i=1}^{p} f_i x_i^2 - \bar{X}^2$$

VI. الانحراف $S = \sqrt{V}$

VII. الوسط الحسابي للانحرافات المطلقة e_m

$$e_m = \sum_{i=1}^{p} f_i \left|x_i - \bar{X}\right|$$

Prof Mustapha
KHALDI

VIII. خواص التباين والانحراف المعياري

1) الخاصية الأساسية

الدالة $g(a) = \frac{1}{N}\sum_{i=1}^{p} n_i (x_i - a)^2$ تقبل قيمة حدية صغرى لما $a = \bar{X}$ و هذه القيمة هي التباين V

2) التغيير التآلفي

إذا كانت A (x_i, n_i) سلسلة إحصائية تباينها V_x و انحرافها S_x و B (y_i, n_i) سلسلة إحصائية تباينها V_y و انحرافها S_y بنفس التكرار و $\boxed{y_i = ax_i = b}$ مع $a \in \mathbb{R}^*$ و $b \in \mathbb{R}$ و من أجل $i \in \{1, 2, 3, \dots, p\}$ يكون لدينا:

$\boxed{V_y = a^2 V_x}$ و $\boxed{S_y = |a| S_x}$

IX. تلخيص سلسلة إحصائية

- يتم تلخيص سلسلة إحصائية بمؤشرين (مؤشر موقع ومؤشر تشتت)
- عموما نختار الثنائية (الوسيط ، الوسط الحسابي للانحرافات المطلقة) : $(Med \text{ ، } e_m)$

  أو الثنائية (الوسط الحسابي ، الانحراف المعياري) : $(\bar{X} \text{ ، } S_x)$
- يمكن استعمال ثنائيات أخرى لتلخيص سلسلة كالثنائية (الوسيط ، المدى) لكن يعاب على المدى تأثره بالقيم الشاذة
- تستخدم أحيانا الثنائية (الوسيط ، الانحراف الربعي) لتلخيص السلسلة و هي ثنائية لا تتأثر بالقيم الشاذة
- تلخيص سلسلة إحصائية يمكننا من مقارنته بسلسلة أخرى ملخصة بنفس الثنائية