I.5.2 – مقاييس التشتت:

تشتت بيانات ظاهرة ما يقصد به مقدار التفاوت أو الاختلاف بين عناصر هذه الظاهرة، وتعتبر بيانات الظاهرة متجانسة عندما تكون قيمها متقاربة من بعضها البعض ونقول في هذه الحالة أنها غير مشتتة. أما إذا كانت بيانات الظاهرة متباعدة وغير متجانسة فنقول أن مفردات الظاهرة مشتتة وغير مركزة. ويقاس تشتت (مدى تباعد أو تقارب) البيانات من بعضها البعض أو من متوسطها بعدة مقاييس منها:

- **المدى R** (Range):

هو الفرق بين أكبر وأصغر قيمة في السلسلة الإحصائية

– **في حالة بيانات غير مبوبة**: يحسب بالعلاقة التالية:

$$R = X_{Max} - X_{Min}$$

– **في حالة بيانات مبوبة:**

يحسب المدى بالعلاقة التالية:

$$R = X_{Max} - X_{Min}$$

❖ في حالة بيانات مبوبة في شكل فئات، المدى هوالفرق بين الحد الأعلى للفئة الأخيرة والحد الأدنى للفئة الأولى.

- **التباين V_X** (Variance): عبارة عن متوسط مربعات انحرافات القيم عن وسطها الحسابي (الفروق بين قيم المتغير الإحصائي ومتوسطها الحسابي). ويرمز له أيضا بالرمز σ^2.

– **في حالة بيانات غير مبوبة**: يحسب بالعلاقة التالية:

$$V_X = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}$$

أو بالطريقة المختصرة كمايلي :

$$V_X = \frac{\sum_{i=1}^{n} X_i^2}{n} - (\bar{X})^2$$

– **في حالة بيانات مبوبة:**

أ– بيانات متقطعة: يحسب بالعلاقة التالية:

$$V_X = \frac{\sum_{i=1}^{k} n_i(X_i-\bar{X})^2}{\sum_{i=1}^{k} n_i}$$ ، الصيغة العادية.

$$V_X = \frac{\sum_{i=1}^{k} n_i X_i^2}{\sum_{i=1}^{k} n_i} - (\bar{X})^2$$ ، الصيغة المختصرة.

مثال: بالرجوع **للمثال 01** ، أحسب التباين بطريقتين مختلفتين.

الحل:

لدينا $V_X = \frac{\sum_{i=1}^{k} n_i (X_i - \bar{X})^2}{\sum_{i=1}^{k} n_i}$

$$= \frac{2\times(1-4)^2+16\times(2-4)^2+16\times(3-4)^2+26\times(4-4)^2+12\times(5-4)^2+8\times(6-4)^2+6\times(7-4)^2+4\times(8-4)^2}{100} = 2.6$$

إذن: مقدار التشتت في هذه البيانات هو: **2.6**.

ب- بيانات مستمرة: في حالة المتغيرمستمر فإننا نعوض X_i بمراكزالفئات x_i في العلاقات السابقة.

مثال: بالرجوع **للمثال 02** ، أحسب التباين بطريقتين مختلفتين.

الحل:

لدينا $V_x = \frac{\sum_{i=1}^{k} n_i (x_i - \bar{X})^2}{\sum_{i=1}^{k} n_i}$

$$= \frac{14\times(7-12.34)^2+7\times(9-12.34)^2+13\times(11-12.34)^2+6\times(13-12.34)^2+10\times(15-12.34)^2+4\times(17-12.34)^2+10\times(19-12.34)^2}{64} = 17.34$$

إذن: مقدار التشتت في هذه البيانات هو: **17.34**.

- **الإنحراف المعياري σ** (Standard Deviation): من أهم مقاييس التشتت وأكثرها استخداما في النظريات والقوانين الإحصائية، يستعمل في المقارنة بين الظواهر، حيث يعتمد عليه كمقياس لقياس تشتت الظواهر في حالة **تساوي المتوسطات الحسابية** لأنه يتأثر بها، وهو الجذر التربيعي للتباين : $\sigma = \sqrt{V_X}$

ملاحظة: لا يمكن اعتماد الإنحراف المعياري للمقارنة بين تشتت توزيعين إحصائيين من نفس النوعية ولهما متوسطات حسابية مختلفة.

مثال: بالرجوع **للمثال 01** ، الإنحراف المعياري هو: $\sigma = \sqrt{V_X} = \sqrt{2.6} = 1.6$

مثال: بالرجوع **للمثال 02** ، الإنحراف المعياري هو: $\sigma = \sqrt{V_X} = \sqrt{17.34} = 4.16$

- **معامل الإختلاف CV** (Coefficient of Variation):

مقياس من أهم مقاييس التشتت النسبي التي تعتمد على فكرة مقارنة البيانات والظواهر في شكل نسبة مئوية اعتمادا على متوسطاتها، والبتالي يسمح بقياس ومقارنة تشتت ظاهرتين ليس لهما نفس المتوسط الحسابي. يعطى **CV** بالعلاقة التالية:

$$CV = \frac{\sigma}{\bar{X}} \times 100.$$

I.2.6- مقاييس الشكل (shape characteristics):

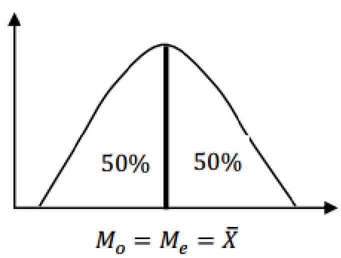
تسمح مقاييس النزعة المركزية ومقاييس التشتت بإعطاء فكرة سريعة عن خصائص توزيع البيانات (تجانسها أو تشتتها)، لكنها غير كافية لوصف البيانات وإجراء المقارنات بين التوزيعات التكرارية، لأنه قد يتساوى توزيعان في المتوسطات والإنحرافات المعيارية ومع ذلك نجدهما مختلفين تماما من حيث شكل التوزيع، الأمر الذي أدى إلى البحث عن مقاييس أخرى أكثر دقة وتفصيل تبين لنا شكل التوزيع وطريقة انتشار بياناته؛ هذه المقاييس تعرف **بمقاييس الشكل** أو مقاييس **الإلتواء والتفرطح**.

أولا: أشكال المنحنيات التكرارية:

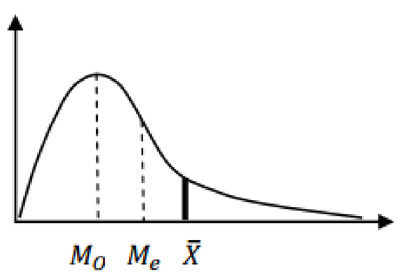
1/ أشكال الإلتواء والتناظر (Skewness and Symmetry):

أ- توزيع متناظر(متماثل):

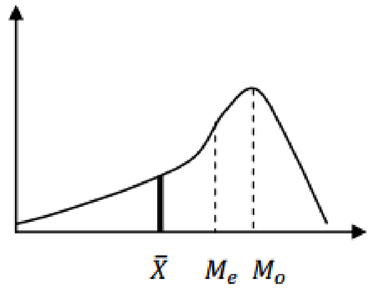
وتتوزع فيه البيانات بشكل متناظر بالنسبة الى الوسيط (المساحة على يمين M_e تساوي المساحة على يساره)، حيث تأخذ المتوسطات الثلاث العلاقة : ($M_o = M_e = \bar{X}$).



ب- التوزيع موجب الالتواء(مائل نحو اليمين): وتتركز فيه البيانات في الفئات الصغرى من التوزيع (المساحة على يمين $\bar{X}$ أقل من المساحة على يساره)، حيث تأخذ المتوسطات الثلاث العلاقة: ($M_o < M_e < \bar{X}$).



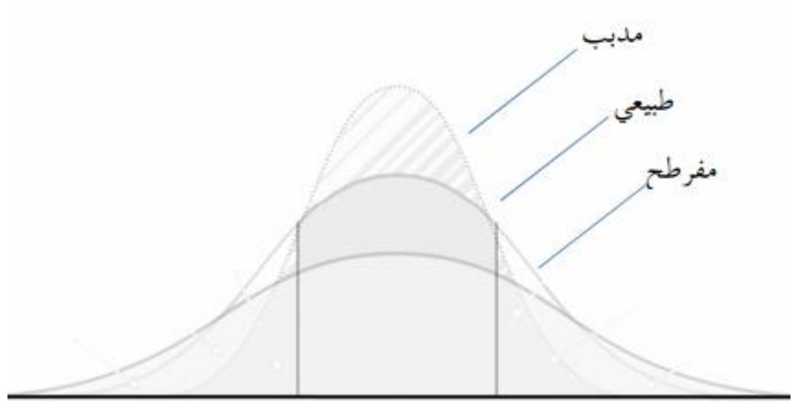
ج- التوزيع سالب الالتواء(مائل نحو اليسار): وتتركز فيه البيانات في الفئات الكبرى من التوزيع (المساحة على يسار $\bar{X}$ أقل من المساحة على يمينه)، حيث تأخذ المتوسطات الثلاث العلاقة: ($\bar{X} < M_e < M_o$).



ملاحظة: إذا كان منحنى التوزيع التكراري ملتويا إلتواء قليل فإن العلاقة التالية صحيحة: $\bar{X} - M_o = 3(\bar{X} - M_e)$.

2/ أشكال التفرطح والتطاول والتوزيع الطبيعي (Kurtosis, Leptokurtic and Normal Distribution) :

يأخذ منحنى التوزيع عدة أشكال؛ فإذا تركزت القيم بشكل كبير في المنتصف (تشتت البيانات ضعيف جدا) يكون المنحنى متطاول (مدبب)، أما إذا تركزت القيم في الأطراف (تشتت البيانات كبير جدا) كان المنحنى مفرطحا أو منبسطا. الأشكال الموالية تبين تطاول وتفرطح المنحنى مقارنة بمنحنى التوزيع المعتدل.



ملاحظة: في أغلب الأحيان لايكون شكل منحنى التوزيع التكراري واضحا وبالتالي يصعب معرفة نوع التوزيع (متناظر، ملتوي، مفرطح، متطاول أو طبيعي) بالعين المجردة، وعليه لابد من مقاييس علمية حسابية تسمح بمعرفة شكل منحنى التوزيع بشكل دقيق.

ثانيا: مقاييس الإلتواء:

1– معامل فيشر α_F:

– **في حالة بيانات غير مبوبة:** يحسب بالعلاقة التالية: $\alpha_F = \frac{\mu_3}{\sigma^3}$

حيث: $\mu_3 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^3}{n}$ و σ : يمثل الإنحراف المعياري.

- في حالة بيانات مبوبة:

يحسب بالعلاقة التالية: $\boldsymbol{\alpha_F} = \frac{\mu_3}{\sigma^3}$

حيث: $\mu_3 = \frac{\sum_{i=1}^{k} n_i (X_i - \bar{X})^3}{\mathbf{n}}$ و σ : يمثل الإنحراف المعياري.

- ❖ في حالة بيانات مبوبة في شكل فئات، فإننا نعوض $\boldsymbol{X_i}$ بمراكزالفئات $\boldsymbol{x_i}$ في العلاقات السابقة.

ويمكن من خلال الإشارة التي يأخذها **معامل فيشر** معرفة شكل الالتواء، كمايلي:

- إذا كان $\boldsymbol{\alpha_F} = 0$ ، فإن منحنى التوزيع متماثل.
- إذا كان $\boldsymbol{\alpha_F} > 0$ ، فإن منحنى التوزيع ملتوي جهة اليمين.
- إذا كان $\boldsymbol{\alpha_F} < 0$ ، فإن منحنى التوزيع ملتوي جهة اليسار.

2- معامل بيرسون P: يحسب بالعلاقة التالية:

$$\boldsymbol{P} = \frac{3(\bar{X} - M_e)}{\sigma} \quad \text{أو} \quad \boldsymbol{P} = \frac{(\bar{X} - M_o)}{\sigma}$$

يحتمل **معامل بيرسون** ثلاث حالات:

- إذا كان: **P=0** هذا يعني أن: (المتوسط الحسابي = المنوال و المتوسط الحسابي = الوسيط)، وهذا يدل على أن منحنى التوزيع **متماثل (متناظر)**.

- إذا كان: $\mathbf{P} > 0$ هذا يعني أن: (المتوسط الحسابي >المنوال و المتوسط الحسابي > الوسيط)، وهذا يدل على أن منحنى التوزيع **ملتوي جهة اليمين**.
- إذا كان: $\mathbf{P} < 0$ هذا يعني أن: (المتوسط الحسابي < المنوال و المتوسط الحسابي < الوسيط)، وهذا يدل على أن منحنى التوزيع **ملتوي جهة اليسار**.

ثالثا: مقاييس التفرطح: هناك عدة مقاييس أهمها:

1- معامل فيشر $\boldsymbol{\beta_F}$:

- في حالة بيانات غير مبوبة: يحسب بالعلاقة التالية: $\boldsymbol{\beta_F} = \frac{\mu_4}{\sigma^4} - 3$

حيث: $\mu_4 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^4}{\mathbf{n}}$ و σ : يمثل الإنحراف المعياري.

– في حالة بيانات مبوبة:

يحسب بالعلاقة التالية: $\boldsymbol{\beta_F} = \frac{\mu_4}{\sigma^4} - 3$

حيث: $\mu_4 = \frac{\sum_{i=1}^{k} n_i (X_i - \bar{X})^4}{\mathbf{n}}$ و σ : يمثل الإنحراف المعياري.

في حالة بيانات مبوبة في شكل فئات، فإننا نعوض $\boldsymbol{X_i}$ بمراكزالفئات $\boldsymbol{x_i}$ في العلاقات السابقة.

ويمكن من خلال الإشارة التي يأخذها المعامل معرفة شكل التوزيع: مفرطح، معتدل أو مدبب، كمايلي:

–إذا كان: $\boldsymbol{\beta_F} = 0$ فهذا يعني أن منحنى التوزيع **متماثل**.

–إذا كان: $\boldsymbol{\beta_F} > 0$ فهذا يعني أن منحنى التوزيع **تطاول**.

–إذا كان: $\boldsymbol{\beta_F} < 0$ فهذا يعني أن منحنى التوزيع **مفرطح**.