جامعة البعث

كلية التمريض

المعهد التقني لخدمات الطوارئ

المحاضرة الرابعة

الاحصاء الحيوي

الدكتور ياسر العمر

المقاييس الرقمية (٢)
NUMERICAL MEASURES (2)

مقاييس التشتت (الانتشار)

Spread Measures: Measures of Dispersion Spread

وللأسباب التي ذكرت في مستهل المحاضرة الثالثة كان من الجدير بنا البحث عن مقاييس أخرى إلى جانب مقاييس النزعة المركزية التي تقيس لنا تشتت قيم الظاهرة عن متوسطاتها وذلك من أجل فهم تفسير ظاهرة الدراسة فهماً جيداً .

هناك العديد من مقاييس انتشار البيانات وكل منها له مفهوم يختلف عن الآخر . ولما كان التشتت لمجموعة من المشاهدات تعني الاختلاف في قيمها . كانت مقاييس التشتت تعطينا فكرة عن درجة الاختلاف والتباعد في قيم مجموعة من المشاهدات . وأهم مقاييس التشتت هي :

١- المدى The Range :

يعرف المدى بأنه الفرق بين أعلى قيمة وأصغر قيمة في مجموعة المشاهدات .

وطريقة حساب المدى كما هو واضح سهلة . لكن هذا المقياس يعتبر ضعيفاً لقياس درجة التشتت إذ إنه لا يؤخذ بالحسبان سوى قيمتين من قيم مجموعة المشاهدات .

مثال : قيم تعداد كريات الدم الحمراء في بعض المرضى:

٧٠٠٠-٨٥٠٠-٩٠٠٠-١٢٠٠٠-١١٠٠٠

فالمدى الأدنى ٧٠٠٠ و المدى الأعظمي ١٢٠٠٠

وبهذا نجد أن المدى في مجموعة المشاهدات في المثال السابق تتراوح بين ٧٠٠٠ – إلى ١٢٠٠٠ .

٢- التباين The Variance

يحدد الفرق بحساب انحراف أو تباعد كل قيمة من المشاهدات عن الوسط الحسابي . وكلما كانت قيمة المشاهدات قريبة من الوسط الحسابي كانت قيمة التباين أو الفرق صغيرة .

فإذا رمز لعدد المشاهدات بـ n ولكل قيمة من المشاهدات بـ × وللوسط الحسابي بـ $\overline{X}$ وللفرق بـ S^2 فإن التباين أو الفرق يحسب كما يأتي :

التباين = مجموع مربع انحراف القيم عن الوسط الحسابي
عدد المشاهدات – 1

$$S^2 = \frac{\Sigma(\times - \times^-)^2}{n-1}$$

مثال : لنعد إلى مثال قيم البوتاسيوم في بلاسما الدم كما يأتي (مقدرة بالمول /لتر) لأربعة عشر عينة.

4.37 , 4.87 , 4.35 , 3.92 , 4.68 , 4.54 , 5.24 , 4.27 , 4.59 , 4.86 , 4.40 , 4.73 , 4.83 , 4.21

$$\bar{X} = \frac{63.96}{14} = 4.57 mmol/L$$

$$S^2 = \frac{(4.37-4.57)^2 + ... + (4.21-4.57)^2}{14-1}$$

$$S^2 = \frac{1.32297}{13} = 0.10177 (mmol/L)^2$$

٣- الانحراف المعياري The Standard Deviation :

يرمز للانحراف المعياري بالرمز SD (Standard deviation) ويرمز له غالباً بالرمز σ (سيغما) وهو الحرف الإغريقي الصغير لحرف (سيغما) الكبير Σ وهو ببساطة عبارة عن الجذر التربيعية المربعة للتباين حيث نأخذ جذره التربيعي ليعطينا الانحراف المعياري .

ويمكن أن تعتبر الانحراف المعياري معدلاً لانحرافات قيم المشاهدات عن الوسط الحسابي . ويرمز عموماً بالرمز (S) في العينة والرمز σ للانحراف المعياري في كامل المجتمع.

وهكذا يحسب الانحراف المعياري رياضياً كما يأتي :

$$S = \sqrt{\text{التباين}} \quad = S^2$$

$$S = \sqrt{\frac{\Sigma(\times i - \times^-)^2}{n}}$$

والجدير بالذكر أن الانحراف المعياري لا يعطي صورة لمتوسط انحرافات القيم عن الوسط الحسابي بل يجذر متوسط مربع الانحرافات عن الوسط الحسابي ويعتبر الانحراف المعياري من أهم مقاييس التشتت لكونه يدخل في معظم عمليات التحليل الإحصائي .

ويمكن أن نحسب الانحراف المعياري بسهولة باستخدام الآلات الحاسبة . ويجب أن نميز أن هناك مفتاحين في بعض الآلات الحاسبة واحداً يرمز له بالرمز S وهو لحساب الانحراف المعياري ضمن العينة ومفتاحاً آخر يرمز له بالرمز σ وهو لحساب الانحراف المعياري ضمن المجتمع المدروس كاملاً .

وفي بعض الآلات الحاسبة نجد أنه يرمز بالرمز σ_n لحساب الانحراف المعياري في عينة ما.

مثال :

في مثال بيانات موجودات البوتاسيوم في بلازما الدم لأربعة عشر كلباً الذي ذكرناه سابقاً وقدرنا فيه قيمة الوسط الحسابي وقيمة التباين أو الفرق يمكننا أيضاً أن نحسب قيمة الانحراف المعياري .

$$S = \sqrt{0.010177} = 0.319 mmal / L$$

ويمكننا القول إن الانحراف المعياري هو أفضل مقياس لدراسة تشتت البيانات لأن الانحراف المعياري تستخدم فيه كافة قيم المشاهدات المدروسة وهو يستخدم لدراسة حد الثقة (والذي سوف ندرسه في الحاضرات اللاحقة) لمعرفة دقة النتائج التي حصلنا عليها في اختباراتنا المعملية . فعلى سبيل المثال لحساب مستوى الثقة ٩٥ % لقيم المشاهدات (على سبيل المثال : البوتاسيوم في بلاسما الدم) نجري الآتي :

الوسط الحسابي ∓ الانحراف المعياري .

$$\text{CI} \quad \%95 = \bar{X} + 2SD$$

حيث أن:

CI= هو حد الثقة (مستوى الثقة) .

$\overline{X}$ = الوسط الحسابي .

S = الانحراف المعياري .

4- معامل الاختلاف Coefficient of Variation :

يعبر عن الانحراف المعياري في بعض الأحيان بنسبة مئوية للوسط الحسابي ولذلك يطلق هذا الاسم على الانحراف المعياري النسبي ويستعمل في التحليل الإحصائي للمقارنة بين تشتت توزيعات مختلفة . وهو ذو أهمية خاصة لمقارنة شدة الإصابة عند استخدام اختبار الاليزا (اختبار المقايسة المناعية الإنظيمية). ويدعى هذا القياس بمعامل الاختلاف ويرمز له بالرمز CV . ويعتبر الانحراف المعياري مقياساً جيداً للتشتت لمجموعة من المشاهدات ولكننا لا نستطيع مقارنة الانحراف المعياري لمجموعتين أو أكثر من المشاهدات لسببين رئيسيين هما :

1- الاختلاف في وحدة القياس .

2- الاختلاف في قيمة الوسط الحسابي .

ولذلك ولمقارنة درجة التشتت في مجموعتين من المشاهدات لا بد من إيجاد مقياس يعتمد على نسبة التشتت بدلاً من القيمة المطلقة للتشتت وهذا المقياس كما ذكر في بداية حديثنا عن معامل الاختلاف الذي يمثل النسبة المئوية للانحراف المعياري من الوسط الحسابي ويحسب رياضياً كما يأتي :

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعياري} \times 100}{\text{الوسط الحسابي}}$$

$$CV = \frac{SD}{\overline{X}} \text{X} 100$$

$$\text{CV} = \frac{\sigma}{\overline{\text{X}}} \times 100$$

مع أن :

CV = معامل الاختلاف.

SD أو σ = الانحراف المعياري.

$\overline{X}$: الوسط الحسابي .

٥- المدى المرجعي The Reference Range :

يهتم في بعض الأحيان بوصف قيم المدى الخاصة المتغير والتي تعرف صحة الأفراد المدروسة وهذا ما يدعى بالمدى المرجعي reference range أو ما يدعى بالفاصل أو الحد المرجعي reference interval ويستخدم هذا المدى لتحديد ما إذا كانت الأفراد يمكن وصفها بأنها تنتمي إلى مجموعات سليمة صحياً أو إلى أفراد مريضة ويدعى المدى المرجعي أحياناً بالمدى الطبيعي normal range ومن الأفضل أن نتجنب استخدام العبارة السابقة لأنها سوف تخلق تشويشاً للقارئ فبينما تشير كلمة طبيعي إلى الأفراد السليمة تصف كلمة طبيعي في التعبير الإحصائي توزيعاً معيناً للبيانات (والذي سيشرح لاحقاً) . و سوف نميز هاتين العبارتين باستخدام أحرف أبجدية إنكليزية صغيرة للحالة الصحية وأخرى أبجدية كبيرة بالنسبة للتوزيع الإحصائي .

على كل إن المدى المرجعي يمكن أن يحسب لحدود الثقة 90 % و 95 % و 99 % إلا أننا غالباً ما نستخدم حد الثقة 95 % وهنا يكون مدى الخطأ 1.96 وبالتالي يحسب المدى المرجعي يحسب كما يأتي :

$$RR = \overline{X} \mp 1.96\ SD$$

مع أن :

$\overline{X}$: الوسط الحسابي ، SD : الانحراف المعياري ويستخدم هنا عندما يكون توزيع البيانات طبيعياً Normal Distribution وعلى كل عندما يكون توزيع البيانات غير طبيعي يمكن أن نحسب المدى المرجعي بفاصل (2.5) بدلاً من 1.96 و 97.5 % بدلاً من 95 % كنسب مئوية من التوزيع التجريبي لقيم المشاهدات .