المحاضرة الأولى

مقرر: الإحصاء الحيوي

كلية التمريض

الدكتور ياسر العمر

المدخل: مفهوم علم الإحصاء

INTRODUCTION: WHAT IS STATISTICS

مفهوم علم الإحصاء What is statistics

يعرف الإحصاء بأنه علم من العلوم الرياضية لأن أسسه وقوانينه مستمدة من القوانين الرياضية البحتة إلا أنه يتميز بكونه علماً تطبيقياً يتناول بالدراسة الناحية الكمية لمختلف العلوم التطبيقية ومنها العلوم الطبية.

ومن هنا يمكننا أن نعرف علم الإحصاء بأنه علم يشمل مجموعة الطرائق والأدوات العلمية التي تتبع في الحصول على المعلومات الكمية والنوعية وتبويبها وعرضها وتحليلها وتلخيصها ولما كانت المعلومات تتحول بالمعنى الرياضي إلى أرقام كمية تتمثل بالبيانات عُرِّف علم الإحصاء بأنه دراسة جمع البيانات وتصنيفها وتبويبها وتلخيصها وعرضها وتحليلها لاستنباط الاستنتاجات العلمية الشاملة من هذه البيانات الجزئية .

إن الأعداد الهائلة للنصوص الأولية في موضوع الإحصاء الأكاديمي تشير إلى أهمية مواضيع الإحصاء الأكاديمي للعلوم الحيوية طلاب ومَنْ يقومون في مجال البحوث العلمية.

وفي الحقيقة فإن هناك كثير من النصوص تتطلب منا أن نكتشف الطريقة الذكية لتمثيل البيانات المطلوبة .

إن المشكلة التي تواجهنا في الإحصاء الحيوي هي كما يأتي :

عندما نقوم بترتيب وتنسيق المشاهدات المعبرة عنها رقمياً في علوم الأحياء فإنك ستجد القيم المحصول عليها على شكل بيانات مشتتة التوزيع (وهذا يحتم علينا معرفة) ما إذا كانت الاختلافات في هذا التوزيع ناجمة عن أسباب متعلقة بموضوع الدراسة أو أنها تكون نتيجة لفرق طبيعي متعلق بخلفيات غاية مواضيع الدراسة وعلاقة ذلك بالعوامل الأخرى ممكناً . كما أننا نحتاج

إلى أن نقيم الوسطَ الحسابي الرقمي فعلياً وإلى أن نمثل هذه الأرقام في طريقة تجعل ارتباط تفسيرها بالعوامل الأخرى ممكناً .

وإذا ما أخذنا كلمة الإحصاء بمعناها الحرفي فإننا نجد أن لها مرادفات عدة فقد ورد ذكرها في القرآن الكريم بمعنى الحصر والعد (وأحصى كل شيء عدداً). وكثيراً ما نصادف العديد من البيانات الإحصائية التي تصدر عن وزارة الصحة والتي تعبر عن أعداد المواليد والوفيات والإصابة بالأمراض المرتبطة بالمسوحات لأمراض الإجهاض وغيرها... ومن هنا كان لفظ الإحصاء دالاً على الطرائق الإحصائية المختلفة التي تستخدم في تحليل وتفسير البيانات الكمية أو العددية . ويعتمد الإحصاء على كثير من النظريات الإحصائية والرياضية والمعادلات المختلفة التي يضعها المتخصصون في العلوم الإحصائية الرياضية.

إن علم الإحصاء يركز على دراسة النقاط الآتية :

١- تصميم الدراسة تصميماً تتضح به المعلومات والبيانات المجموعة اتضاحاً فعالاً.

٢- جمع البيانات .

٣- تحليل البيانات .

٤- تمثيل المعلومات وتلخيصها تلخيصاً مناسباً وذلك باستخدام الأشكال أو الجداول المناسبة .

٥- تفسير التحليل للنتائج بطريقة تربط بين موجودات الدراسة بدقة .

إن عرض النقاط السابقة ودراستها في علم الأحياء يُعبر عنه بالقياس الحيوي (biometry) لكننا نؤثر استخدام تعبير أشيع وهو ما يدعى بعبارة الإحصاء Statistics لنغطي كافة نقاط الدراسة .

وبهذا يتبين أن الإحصاء قد أصبح إحدى الوسائل الأساسية في علوم الأحياء الحديثة .

دور الإحصاء في تسلسل خطوات البحث العلمي

The Role of Statistics in the Steps of Scientific Research

من التساؤلات الأساسية الشائعة لدى كلٍّ من طلبة العلوم العلوم الطبية هو: لِمَ أحتاج إلى أن أدرس الإحصاء ؟

إن الأساس الرياضي البحث للإجابة على ذلك وتعطي تفسيرات غير مؤكدة غالباً . وفي الحياة المهنية على العموم كثير من الأمثلة التي توضح أهمية الإحصاء في العلوم الحيوية وغيرها من العلوم:

١- المنشورات العلمية المطبوعة The Published – Scientific Literature إن مثل هذه المنشورات مليئة بالطرائق الإحصائية المستخدمة والتي تظهر بوضوح استخدام تعبير (الوسط الحسابي ∓ الوسط الحسابي للخطأ للمعياري) أي [mean ∓ Standard error of mean (SEM)] للحصول على المعنوية الإحصائية بواسطة قيم الاحتمالية (P) أو باستخدام اختبارات t ستدنت (student t) أو تحليل مربع كاي (χ^2) أو استخدام طرق تحليل التباين . ونجد أن المعلومات الممثلة فيها مختصرة ومبوبة . وبدون إلمام الدراسة بالتحليل الإحصائي لا نقبل استنتاجات عمل هذا الباحث مطلقاً ، لأننا نكون عندئذٍ غير قادرين على فحص قوة الدعم والربط للبيانات مع الواقع . باختراع الحاسب الذي يقوم بتنسيق وتنظيم البيانات والكثير من حسابات حول المشاهدات التي حصلنا عليها ويقوم أيضاً بتلخيصها وتبويبها وتبيان أهميتها وحتى نقوم بكل ما ذكرناه في جهاز الحاسب ليُحقق الفائدة المرجوة من هذه المعلومات فإننا نحتاج حتماً إلى نظرة شاملة بدراستها إحصائياً .

٢- وفي علوم صحة الإنسان: تتزايد أعداد الخدمات التشخيصية المستقلة والتي تحلل العينات من أجل تحقيق الفائدة في رصد الصحة Health monitoring والمحافظة عليها جيداً . هذا العمل يجرى بواسطة الخدمات المخبرية والتي توجب التركيز دائماً على نوعية التحكم والدقة في

القياسات المتبعة للأغراض التشخيصية وكونَنا قادرين على تقديم دلالات واضحة لتفسير النتائج المحصول عليها في المخابر .

٣- الصناعات الدوائية والكيميائية الزراعية Pharmaceutical and Agrichemical Industries : وهي تتطلب إثبات كل من أمان منتجاتها وفعاليتها بطريقة لا تقبل الجدل . مثل هذه البيانات لآلية هذه الصناعات تتطلب طرائق إحصائية لتقييم وتوضيح أساس كل من هاتين النقطتين المدروستين . ففي حالة الصناعات الدوائية نجد أن تطويرها يحتاج إلى فهم أهمية تصميم الدراسة للتأكد بما يكفي من أعداد الأفراد المستخدمة في المجموعات المعالجة كي تجرى تجارب مثل هذه الصناعات بما هو جدير بالثقة. إن لجنة ترخيص المنتجات الطبية يطلب منها فهم علم الإحصاء لتقييم مدى فعالية المواد العلاجية المصنعة وتأثيرها على الانسان أو المريض .

٤- ويزداد التركيز حالياً على أنظمة الأمانة ونوعية الطعام الخاص بالاستهلاك البشري . لأن المنتجات من أصل ي ذات أهمية في هذا المجال وتحديد مدى أمانها يقع على عاتق الطبيب البيطري فعلى سبيل المثال : المنتج الدوائي المتعلق بسحب زمني معين قبل الذبح يعتمد على التداخل الدوائي وديناميكية الدواء الخاصة بالمنتج وهذا يتعلق على سبيل المثال بمنتج الحليب بعد عملية معالجة ال وكذا بالنسبة للثمالات النسيجية الناجمة عن المبيدات العشبية والمبيدات الحشرية وإمكانية تلوث الذبائح بالصادات الحيوية المقاومة للجراثيم . وعلى أية حال فإن التوصيات والمقترحات وكذا القوانين الناظمة المناسبة تُقيّم بالدراسات التجريبية وكذا التقييم الإحصائي . وهذا يتطلب من الخبراء أن يكون لديهم خبرة بالطرائق الإحصائية المناسبة كي يؤدوا عملهم في سن القوانين المناسبة والصحيحة اعتماداً على هذه الطرائق .

ويمكن أن نوجز فوائد علم الإحصاء الأساسية في ثلاثة مجالات متعلقة بالعلوم الطبية:

١- تعتبر دراسة الطرائق الإحصائية واستعمالها في التجارب الحيوية والمشاهدات التجريبية أداة مهمة في جمع وتبويب وتحليل البيانات وتفسير النتائج الحيوية .

٢- إن دراسة العلوم الإحصائية ونظرياتها المختلفة تسهل وتساعد على فهم الباحث في الحصول على الحقائق وفق أفضل النتائج وأدها ومطابقتها فيزيولوجيا لوقائع الأحداث المدروسة .

طريقة البحث الإحصائي ومراحله

Method of Statistical Approach and its Steps

تشمل طريقة البحث الإحصائي مراحل مختلفة تختلف تبعاً لإطار الدراسة والهدف المطلوب إنجازه إلا أنها عموماً تشمل المراحل الآتية :

١- جمع البيانات الإحصائية بالطرائق المناسبة والتي سنتطرق إليها في فصول لاحقة .

٢- تصنيف وتبويب البيانات .

٣- عرض البيانات بطريقة أو أكثر من طرائق العرض المناسبة .

٤- دراسة البيانات المجموعة وتحليلها بالطرائق الإحصائية الوصفية أو التحليلية أو بكلتيهما معاً

٥- استخلاص النتائج المحصول عليها من تحليل بيانات الدراسة .

إلا أنه يجب أن ننوه هنا بأن كل مرحلة من مراحل هذه الدراسة مهمة ويجب إتمامها بدون أخطاء مع اعتبار هذه المراحل مرتبطاً بعضها ببعض وأن كل مرحلة تبدأ من نهاية المرحلة السابقة وتعتمد عليها وعلى دقة نتائجها.

أنواع المتغيرات Types of Variables

المتغير هو عبارة عن قيمة خاصة تتعلق بها صفة أو عضو قابل للتغير في النوع أو الكم من مكرر إلى آخر في نفس المجتمع المدروس ، وقد تكون الصفة المتغيرة وصفية أي لا يمكن

قياسها مباشرة كاللون أو الحالة الصحية للإنسان وقد تكون الصفة المتغيرة صفة كمية كالصفات والخواص التي يمكن قياسها مباشرة مثل الأطوال والأوزان ... الخ .

وهنا يجب أن نتحدث عن القيمة الإحصائية والتي تعبر عن قيمة إحدى صفات المجتمع المدروس المتغير والتي يمكن أن تأخذ قيماً تختلف في حدودها من القيم الفردية في أحد المجتمعات إلى قيم فردية أخرى أو من قيم شاملة لمجموعة من الأفراد إلى مجموعة أخرى . فعلى سبيل المثال صفة الطول والوزن وتعداد الدم وفعالية الأنزيمات ولون غطاء الجلد والنسب المئوية للنساء الحاملة في مجتمع ما .. إلخ ويرمز للقيمة الإحصائية بالرمز (X) فإذا كان لدينا مجموعة أطوال لمجموعة من الأفراد فيأخذ كل فرد طولاً فردياً مثلاً ويرمز له بـ X_1 والثاني بـ X_2 والثالث بـ X_3 .. وهكذا حتى الطول الأخير Xn ويكون مجموع القيم الإحصائية هو :

$$\sum_{1}^{n} X = X_1 + X_2 + X_3 + ... + \times n$$

حيث يعبر Σ عن المجموع المشتق من الحرف الإغريقي Sigma كما يعبر عن المجموع بالرمز (S) والمشتق من الكلمة الإنكليزية Sum أو Summation إلا أن ذلك غير مرغوب فيه منعاً للخلط بين هذا الرمز ونفس الرمز (S) الذي يعبر عنه أحياناً بالانحراف القياسي كأحد مقاييس التشتت. ومن الملاحظ أنه وضع تحت الرمز Σ الرقم ١ وفوقه الحرف n وذلك للتعبير عن أن هذا المجموع هو عبارة عن مجموع الأفراد في مجتمع الدراسة والتي تأخذ القيم من الفرد الأول (١) حتى الرقم الأخير (n) وعلى نحو واضح فإن بعض هذه المميزات المذكورة الخواص يكون أكثر قابلية لتقديرها كمياً من الأخريات . (ومن أجل بعض المتغيرات يمكننا أن نحول الرقم العددي إلى فئة رقمية وبذلك يمكننا إنشاء نوع من المستوى العددي إلا أن بعض القيم الأخرى يمكن أن تكون على شكل مستوى رقمي حقيقي تمتد عندها قيمة المتغير) .

كما أن البيانات الرقمية تأخذ أشكالاً مختلفة بما أن الفهم الطبيعي للبيانات وتصنيف المتغيرات هو الخطوة الهامة الأولى في اختيار الطريقة الإحصائية المناسبة .

ويمكن أن يُميَّز الأنماط الرئيسية للمتغير بطريقة نظامية وذلك بتحديد إمكانية أخذ المتغير لقيمة من قيمتين مميزتين ، أو قيمة واحدة من عدة قيم مميزة ، أو أي قيمة ضمن المجال المعطى لهذا المتغير . وعملياُ يمكن أن يأخذ أحد الأنواع الآتية :

أ- متغير نوعي/أو فئوي Qualitative/ Categorical variable :

وهو عبارة عن عدد ينتمي إلى فئة واحدة من مجموع فئتين اثنتين أو أكثر من الفئات المميزة لهذا المتغير فالمتغير من نوع ثنائي/بيزي (a binary v.) أو ما يدعى dichotomous وهو نوع خاص من المتغيرات من أنماط المتغيرات الفئوية الاسمية والتي تعرف فقط في حال وجود فئتين ، كأن نقول مثلاً :إمرأة حامل أو غير حامل أو أن نقول هذا الفرد ذكر أو أنثى . وتتراوح قيم هذا المستوى من المتغيرات ما بين القيمة الصفر إلى القيمة واحد (٠-١) .

ويمكن أن تلخص المستويات الخاصة بالمتغير الفئوي بما يأتي :

١- مستوى / قياسي / اسمي Nominal Scale : وهي عبارة عن فئات مميزة تعرف المتغير بأنها ذات قيم غير مرتبة Unordered وكل منها يمكن أن يشار إليه باسم كأن نقول لون البشرة الجلد (بني – رمادي) .

٢- مستوى ترتيبي Ordinal Scale : وهو يضم فئات متغير تتميز ببعض الترتيبات الداخلية لكن هذا الترتيب لا يتضمن أي معنى بالنسبة للمتغير وتعرف الفواصل بين الفئات المختلفة حسب المتغير : على سبيل المثال قياس حالة الجسم في مصطلح الأمراض الباطنية Body Condition Score وكذا كثافة الومضان الفلوراسيني للخلايا عند استخدام المجهر الومضاني الفلوراسيني ودرجة حركية النطاف المنوية في العينة . هذه المستويات تعطي غالباً قيماً عددية تتراوح بين ١ إلى n .

ب- المتغير الكمي Quantitative variable

وهو يتضمن قيماً عددية معرفة جيداً بمستواها وهنا نميز نوعين من المستويات أو المتغيرات الكمية :

١- المستوى غير المستمر (متغيرات منقطعة) :

A discrete (Discontinuous) Scale

نقول عن المتغير الإحصائي X إنه متغير منقطع إذا كان $X \in [a , b]$ وكانت هناك قيمة واحدة على الأقل تقع ضمن المجال ويستحيل للمتغير X أن يأخذها. ومثال هذا المستوى ، البيانات التي يمكن أن تأخذ قيماً خاصة ذات تعداد نموذجي مثل تعداد الكريات الحمراء أو البيضاء (أو عدد مرات الحمل) . لنأخذ مثلاً عدد البشر المعالجة يومياً في المشفى ولنفرض أن هذا العدد لا يقل عن ٢ ولا يزيد عن ١٠ ،

أي $X \in [2 , 10]$ وهنا نقول إننا أمام متغير إحصائي منقطع ،أما بالنسبة للعدد ٢ ، ٣ مثلاً فلا يمكن أن يأخذ المتغير X هذه القيمة .

٢- المستوى المستمر أو المتغيرات الإحصائية المستمرة

Variable a Continuous Scale or a Continuous

وتكون هنا جميع القيم مقبولة نظرياً في هذا المستوى وربما تكون القيم ذات حد أعلى وحد أدنى . وبالتالي فإن المتغير X هنا يأخذ أية قيمة من مجموعة الأعداد الحقيقية التي تمثل بمستقيم الأعداد الحقيقية R (real numbers) .

فإذا كان المتغير X متغيراً إحصائياً مستمراً وكان $X \in [a,b]$ فإن X يمكن أن يأخذ أية قيمة من a إلى b بما فيها الأعداد الصحيحة والنسبية والعشرية غير النسبية مثل π والقيمة e . ومثال ذلك صفات الطول والوزن والسرعة وتركيز مكون كيميائي في الدم أو البول . وتكون قيم مستوى هذا المتغير ذات مدى غير محدود أي يمكن أن يصل بقيمته إلى اللانهاية . والمثال الرقمي

لهذا المتغير أن نأخذ كمية قيم تعداد الكريات البيضاء لمجموعة من الأفراد في مدينة من المدن السورية فإذا كانت هذه القيم في حدها الأعظمي في هذه المدينة ٧٥٠٠٠ خلية بالميلليتر وأدنى قيمة لإحدى الأفراد في هذا المجتمع هو ٦٢٠٠ خلية بالميلليتر فإن X يضم المجال $X\in[6200,7500]$ فإذا ما اخترنا أي فرد في هذا المجتمع فإن هذه القيمة يمكن أن يكون واقعاً ضمن المجال المذكور أعلاه مهما كانت دقته كالقيمة ٧٠٠٠ خلية بالملليتر أو القيمة ٧١٠٠ أو القيمة ٧٣٠٠ خلية بالملليتر على سبيل المثال.