

# الإحماء الحياي الأخطاء العيّارية المجتمعات، العينات، الأخطاء العيّارية

Populations, samples, standard errors and confidence intervals المحاضرة رقم ١٣: الخطأ المعياري

الدكتور محمل يونس حجير

أستاذ تقويم الأسنان والفكين بكلية طب الأسنان بجامعة دمشق دكتوراه دولة (PhD) في تقويم الأسنان والفكين من جامعة غلاسكو Glasgow بريطانيا ماجستير في أصول البحث العلمي في بحوث العناية الصحة والاجتماعية من جامعة شيفيلا Sheffield - بريطانيا عضو مجموعة الدراسة البحثية الوجهية القحفية في جامعة غلاسكو Glasgow - بريطانيا محكم دولي في المجلة البريطانية لتقويم الأسنان و مجلة تقويم الأسنان والبحث الوجهي القحفي ومجلة «التطور في تقويم الأسنان»

## الخطأ المعياري Standard Error

#### الخطأ المعياري في الوسط الحسابي

- كم هي جودة «الوسط الحسابي» المحسوب في العينة في تخمين الوسط الحسابي للمجتمع الأم ؟
- لكي نجيب على هذا السؤال يجب أن نحدد مقدار «الشكوكية» uncertainty في الوسط الحسابي المحسوب من العينة الملتقطة

• إذا عدنا إلى إحدى الأمثلة السابقة ، الوزن الولادي في دراسة O'Cathain وزملائه (٢٠٠٢) ، بدراسة ٣٢٢٦ طفل وليد ، تبين أن الوسط الحسابي هو ٣٦.٣ كغ ، والانحراف المعياري ٥٥. • كغ ، وكان التوزع شبه طبيعي.

- افترض أننا أخذنا عينة عشوائية من هذا المجتمع وحسبنا الوسط الحسابي للعينة ، إذا أخذنا عينة عشوائية أخرى وحسبنا الوسط الحسابي سوف نحصل ربما على مقدار آخر (أي تخمين آخر) للوسط الحسابي للمجتمع الأم.
- ولنأخذ مثلا ١٠٠ عينة من نفس الحجم (حجم العينة = ٤)، سوف نحصل على توزع للأوساط الحسابية للعينات المأخوذة ويمكن إظهاره بصريا على هيئة مخطط نقطي، أو منسج

100 Random samples of size 4 8 8 B 8 8 8 100 Random samples of size 16 2000000 88 Sample size 100 Random Samples of size 25 100 Random samples of size 100 3.5 2.5 3.0 4.0 Mean birthweight (kg)

- إن هذه الأوسط الحسابية للعينات الملتقطة تتراوح بين ٤.٢ كغ إلى ٠.٤ كغ ولكن إذا رفعنا حجم العينة = ١٦ فردا ، وأخذنا ٠٠٠ عينة ، ونظرا إلى توزع اعتيان الوسط الحسابي هنا نجد أن المدى يتراوح بين ١.٣ كغ إلى ٣.٧ كغ
- ويعود السبب في ذلك أن الوسط الحسابي لعينات كبيرة يقوم بامتصاص أو تمييع تأثير المشاهدات الكبيرة جدا أو الصغيرة جدا في العينة أكثر مما هو عليه الحال لو كانت العينات المسحوبة بحجم عينة صغير وتمتلك مثل هذه المشاهدات المتطرفة ...

#### العينات

• إن التباينية Variability في لعينات الأوسط الحسابية المشاهدة تعطينا فكرة عن مقدار «الشكوكية» أو «عدم الثقة» المرافقة للتخمين الذي حصلنا عليه للوسط الحسابي للمجتمع الأم.

• حيث نكون بدرجة عالية من «عدم الثقة» عندما يكون حجم العينة صغيرا، ولكن تصبح حالة «عدم الثقة» صغيرة عندما يكون حجم العينة كبيرا ...

88 8 4.0 3.0 3.5 Mean birthweight (kg)

8

2.5

• إن الشكل الذي شاهدناه قبل قليل يوضح كيف أن مقدار التشتت في اعتيان الأوساط الحسابية يقل مقداره كلما كان عدد افراد العينة المسحوبة أكبر.

• وفي الواقع فإن الخصائص التالية هي الخصائص الواسمة لتوزع الأوساط الحسابية ...

#### Properties of the distribution of sample means

The mean of all the sample means will be the same as the population mean.

The standard deviation of all the sample means is known as the standard error (SE) of the mean or SEM.

Given a large enough sample size, the distribution of sample means, will be roughly Normal regardless of the distribution of the variable.

- الوسط الحسابي لتوزع الأوساط الحسابية هو الوسط الحسابي للمجتمع
- الانحراف المعياري لتوزع الأوساط الحسابية هو الخطأ المعياري للوسط الحسابي SEM .
- في حال كانت العينات المسحوبة كبيرة ، فإن توزع الأوساط الحسابية سوف يقارب التوزع الطبيعي بغض النظر عن التوزع المتغير نفسه المدروس ...

• إن الخطأ المعياري (أو التباينية) في التوزع الاعتياني للأوساط الحسابية ، يتم تقييمه من خلال الانحراف المعياري.

• إذا كنا بصدد الحديث عن المجتمع الأم ، ونحن نعلم الانحراف المعياري بالمجتمع الأم  $\sigma$  فإن الخطأ المعياري يعطي بأنه حاصل قسمة  $\sigma$  على الجذر التربيعي لـ n .

• في الحقيقة ، إن الباحث لا يقوم بسحب العينة إلا مرة واحدة فقط ( إلا أنه في بعض الأحيان يمكن أن يسحب عينة اخرى لأهداف تأكيدية) .... وبالتالي فإن الدراسة الأحادية على هذه العينة الوحيدة تقدم لنا وسطا حسابيا متعلقا بالعينة (الإكس بار) ... و هذا هو أفضل «تخمين» لنا (والوحيد المتواجد بين أيدينا) للوسط الحسابي للمجتمع الأم μ.

• كنا أن هذه العينة تقدم لنا انحراف معياريا SD وهو أفضل تخمين متواجد لدينا للتعبير عن الانحراف المعياري بالمجتمع الأم o .

• وبالتالي من خلال دراسة واحدة يستطيع الباحث أن يخمن الانحراف المعياري لتوزع الأوساط من خلال قسمة SD على الجذر التربيعي لعدد الأفراد n بدون أن يضطر إلى إعادة الدراسة مرة أخرى.

مثال :

#### Worked example: Standard error of a mean – birthweight of preterm infants

Simpson (2004) reported the birthweights of 98 infants who were born prematurely, for which n = 98,  $\overline{x} = 1.31$  kg, s = 0.42 kg and  $SE(\overline{x}) = 0.42/\sqrt{98}$  = 0.04 kg.

The standard error provides a measure of the precision of our sample estimate of the population mean birthweight.

#### خصائص الأخطاء المعيارية

• إن الخطأ المعياري هو مقياس «لدقة» و «إحكام» التخمين الآتي من العينة

#### Properties of standard errors

The standard error is a measure of the precision of a sample estimate. It provides a measure of how far from the true value in the population the sample estimate is likely to be. All standard errors have the following interpretation:

- A large standard error indicates that the estimate is imprecise.
- A small standard error indicates that the estimate is precise.
- The standard error is reduced, that is, we obtain a more precise estimate, if the size of the sample is increased.

#### نظرية الحد المركزي Central Limit Theorem

- يمكن ملاحظة أن توزع الأوساط الحسابية يظهر لنا شكل التوزع الطبيعي ، بغض النظر عن التوزع الأصلي للمتغير المدروس بين أفراد العينة ، وكلما زاد حجم العينات المأخوذة لحساب الوسط الحسابي ، كلما قارب هذا التوزع التوزع الطبيعي ...
- من الناحية التقنية ، إن حقيقة أنه بالإمكان تخمين «الخطأ المعياري» من خلال سحب عينة واحدة يعود إلى نظرية نسميها نظرية الحد المركزي Central Limit Theorem
  - وإن هذه الخاصية تسمح لنا بتطبيق التقنيات التي نصفها على عدد كبير من الحالات المختلفة

• لكي نوضح لكم نظرية الحد المركزي ، دعنا نأخذ هذا المثال ... افترض أننا استخدمنا جدول الأرقام العشوائية ..

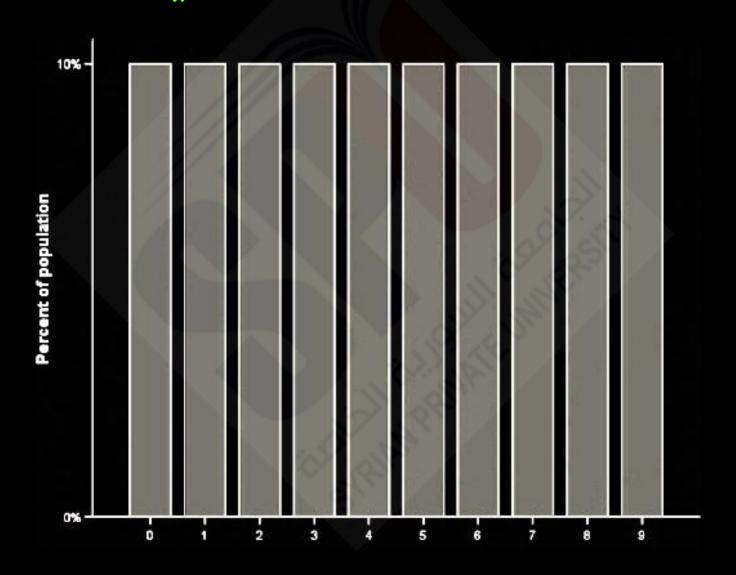
Table T2 Random numbers table

Each digit 0-9 is independent of every other digit and is equally likely to occur.

95188	25011	29947	48896	83408	79684
96046	51589	84509	98162	39162	59469
94731	63786	88290	60990	98407	43473
14846	86619	04238	36182	05294	43791
03876	89100	66895	89468	96684	95491
23327	78957	50987	77876	63960	53986
54870	35009	84524	32309	88815	86792
90047	44763	44534	55425	67170	67937
29068	74625	90665	52747	09364	57491
58563	84810	22446	80149	99676	83102
03778	05031	90146	59031	96758	57420
84334	54827	51955	47256	21387	28456
91577	43019	67511	28527	61750	55267
67970	29162	60224	61042	98324	30425
94071	63090	23901	93268	53316	87773

• لو افترضنا أن المجتمع الأم يحوي هذه الأرقام جميعا من صفر إلى تسعة (ضمن جدول الآرقام العشوائية) ، ولو أردنا أن ننظر إلى توزع هذا المجتمع لوجدنا أن كل رقم يرد بنسبة ١٠ % في هذا المجتمع ...

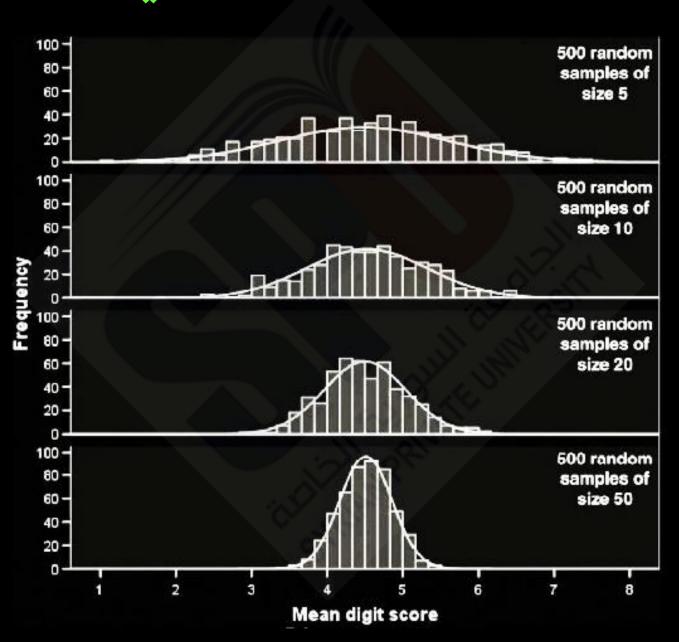
• (ملاحظة: نسمي هذا التوزع باسم التوزع موحد الشكل Uniform Distribution). وإن الوسط الحسابي لهذا المجتمع هو ٥٤٠.



• لنفترض الآن ، أننا أخذنا عينة عشوائية تتالف من خمسة أرقام من هذا التوزع ، وحسبنا الوسط الحسابي ، وأعدنا العملية ، • • مرة ، في كل مرة سوف نحصل على تخمين للوسط الحسابي للمجتمع الأم ..

مثلاً قد نحصل على ٢٠٤، ٣٠، ٣٠ الخ

قيم الوسط الحسابي لكل عينة مسحوبة لا يمكن أن تكون قريبة من الصفر أو من التسعة ، ففي الغالب سوف تدور حول ٥٤٠.



• من الضروري أن نوضح أنه أيا كان التوزع الأم للمتغير المدروس ، فإن توزع الأوسط الحسابية سوف يكون قريبا من التوزع الطبيعي طالما أن العينات المسحوبة بحجم كبير...

• وكلما كان n أكبر في توزع اعتيان الأوساط الحسابية فإن التوزع سوف يأخذ المنحى الطبيعي .

## شكرا جزيلا لحسن استماعكم

myhajeer@gmail.com