الإحصاء الحيوي

التوزعات

Distributions

المحاضرة رقم ١١ : التوزع الطبيعي

الدكتور

محمد يونس حَجير

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 شيفيلد 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 «آنجل التقويمية السنية» ومجلة «التطور في تقويم الأسنان»

احتمالية المتغيرات الحصيلية المستمرة

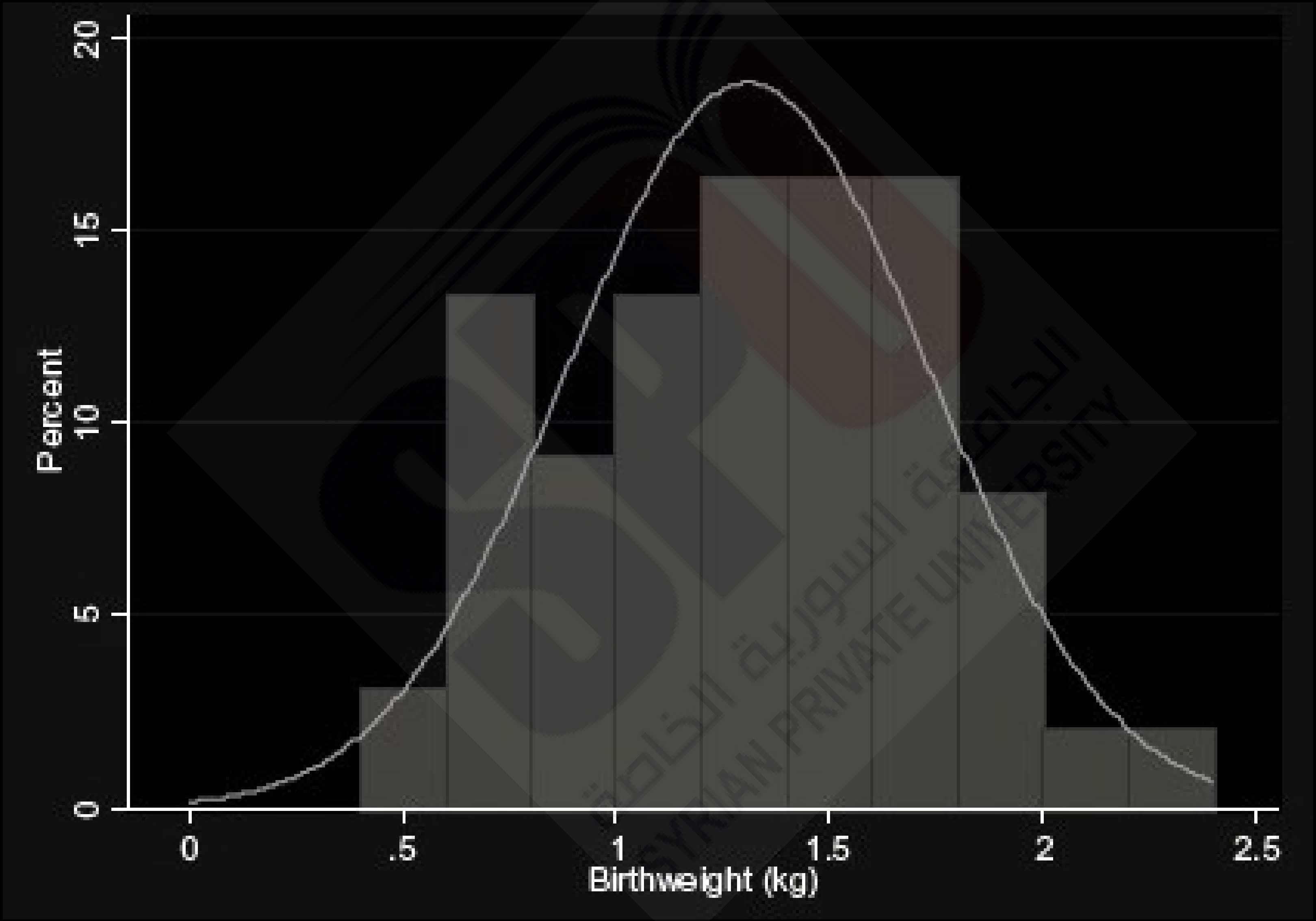
- بالنسبة للمتغيرات الكمية المستمرة ، مثل الوزن الولادي Birth weight ، أو ضغط الدم Blood Pressure فإن مجموعة البيانات المحتملة غير محدود .

- وبالتالي نكون أكثر اهتماما باحتمالية الحصول على قيم بين حدود معينة بدلا من اهتماما باحتمالية الحصول على قيمة مفردة .

- مثلا : ماهي احتمالية الحصول على ضغط دموي انقباضي ١٤٠ ملم زئبق أو أكثر ؟

احتمالية المتغيرات الحصيلية المستمرة

- الشكل التالي يوضح لنا منسجا يتألف من التكرار النسبي على الخط العمودي (محور الـ Y) ، إن ميزة استخدام التكرار النسبي هو أن المقياس العمودي لعدة مناسج أخرى التي تدرس نفس الحصيل ، ولكن بأحجام عينات مختلفة ، سوف يكون هو نفسه .


Percent
0
5
10
15
20
0
.5
1
1.5
2
2.5
Birthweight (kg)
الجامعة السورية الخاصة
SYRIAN PRIVATE UNIVERSITY

احتمالية المتغيرات الحصيلية المستمرة

- إذا ما تخيلنا أننا نمتلك عينة أكبر بكير من تلك المعروضة بالشكل السابق (دراسة سمبسون ٢٠٠٤ على ٩٨ طفل وليد) ، وأخذنا فواصل أصغر وأصغر من فئات الوزن الولادي (فواصل أصغر من ٠.٢ كغ الموجودة في الشكل السابق)..

- وبالتالي فإن المنسج المتشكل سوف يبدأ بالظهور على هيئة منحنى ناعم ، وفي مثل هذه الحالات فإن المشاهدات يمكن مقاربتها على هيئة خط مستبطن منحن ، والمشاهد في الشكل السابق .

احتمالية المتغيرات الحصيلية المستمرة

- نسمي هذا المنحنى : «التوزع الاحتمالي» Probability Distribution ، وهو المكافئ النظري للتوزع التكراري النسبي التجريبي .Empirical Relative Frequency Dis

- إن التوزيعات الاحتمالية تستخدم من أجل حساب احتمالية وقوع القيم المختلفة ، على سبيل المثال : ما هي احتمالية ولادة مولود بوزن ولادي ٢ كغ أو أقل ؟

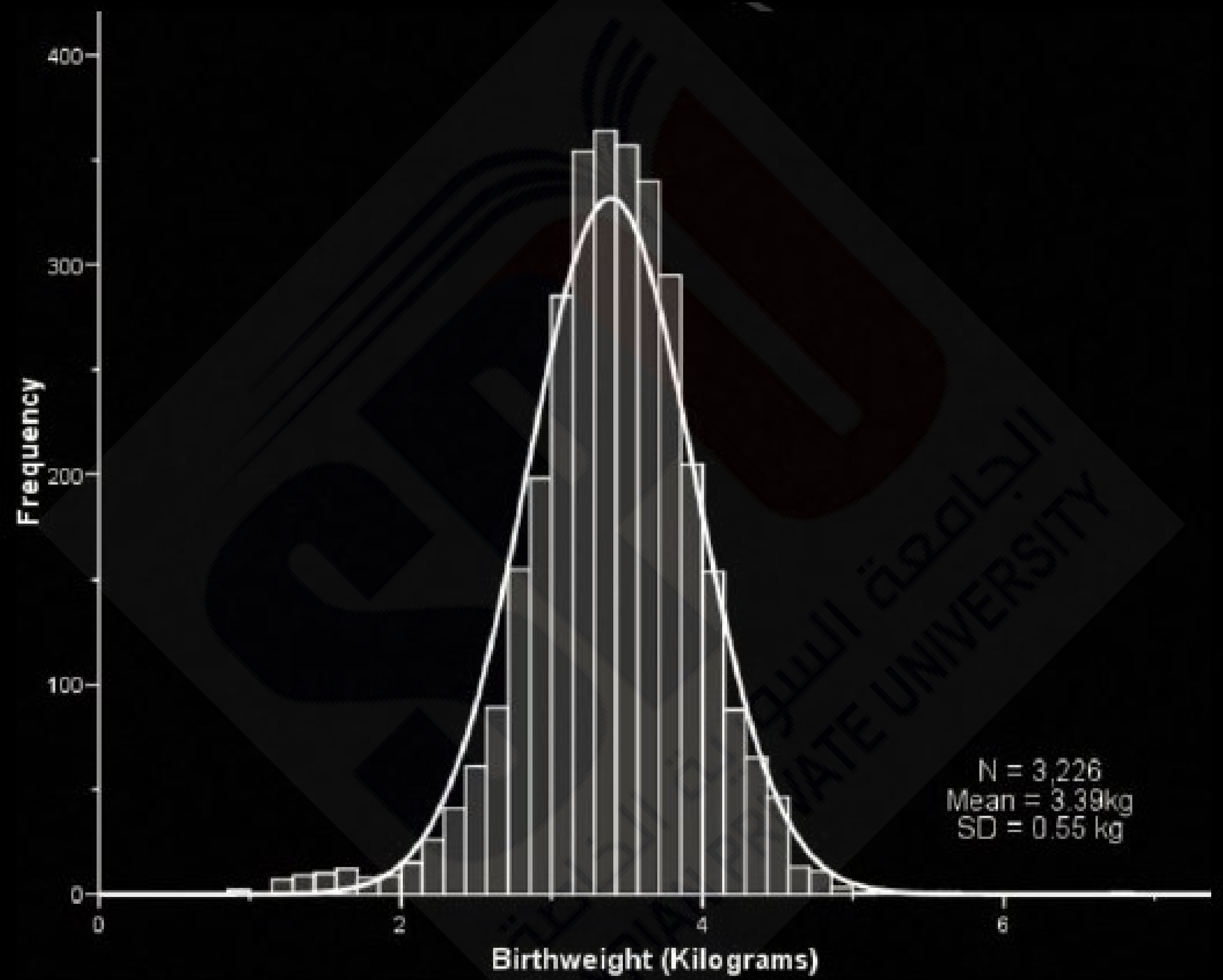
- إنه من الشائع في البيانات الطبية ان المنسج المتولد عن متغير كمي مستمر والذي تم الحصول عليه من أحد القياسات المجراة على مجموعة من المرضى أو الأشخاص المتطوعين سوف يمتلك خصائص التوزع ذي «الشكل الجرسي المقلوب» و المتناظر ...

The Normal Dist. التوزع الطبيعي

- إن التوزع ذا الشكل الجرسي والمتناظر المذكور سابقا نسميه «التوزع الطبيعي» أو Normal Distribution ..

- وهو واحد من أهم التوزعات في علم الإحصاء الحيوي .

- مثال على هذا التوزع الطبيعي ، دراسة «اوكاثين» وآخرين لعام ٢٠٠٢ حول الوزن الولادي المجراة على ٣٢٢٦ طفل حديث الولادة (بالكغ) والمشاهدة بالشكل التالي ...



دراسة O’Cathain وآخرين لعام ٢٠٠٢

التوزع الطبيعي

- وفي كتب الإحصاء الطبي من أجل التمييز بين عبارة normal range أي المدى الطبيعي ، و التوزع الطبيعي Normal distribution ، تم الاتفاق على استخدام الحرف الإنكليزي الكبير **N** في بداية الكلمة عندما نتحدث عن التوزع الطبيعي.

((طبعا هذا الاتفاق لا ينطبق على الكتابة العربية)).

- في المثال المذكور سابقا ، فإن المنسج المعبر عن بيانات العينة هو تخمين للتوزع في المجتمع الأم للوزن الولادي للأطفال حديثي الولادة ، ويمكن تخمين هذا التوزع المجتمعي من خلال «مراكبة» و «إسقاط» هذا المنحنى الناعم ذي الشكل الجرسي المعبر عن «التوزع الطبيعي» على البيانات المجموعة ..

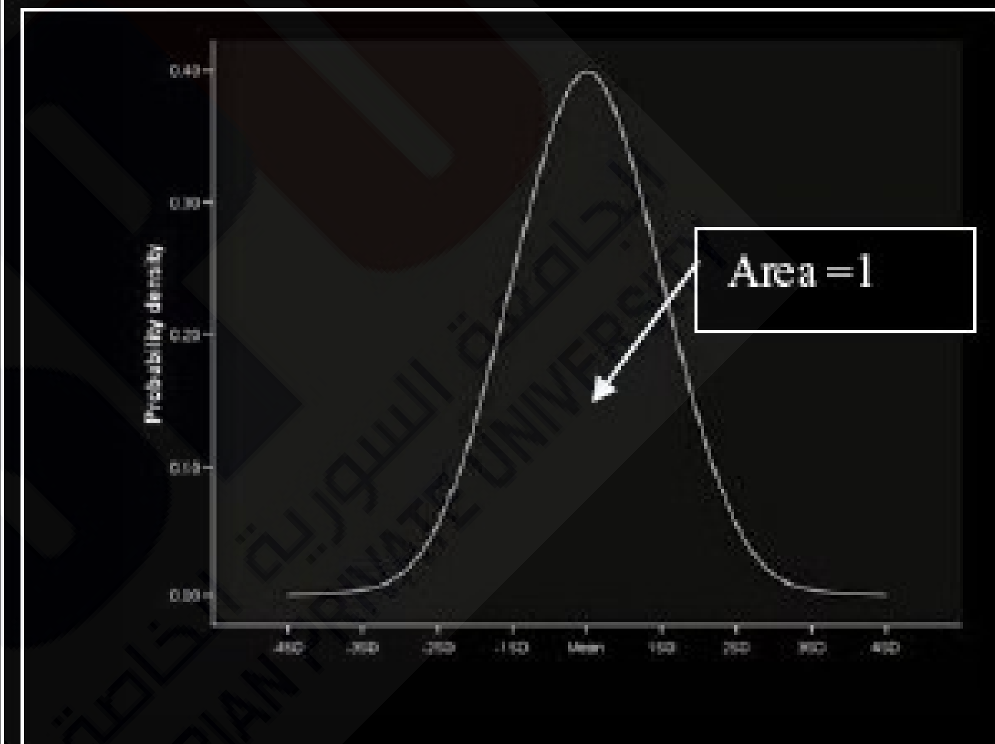
التوزع الطبيعي

- إذا افترضنا أنه كان بالإمكان الاطلاع على الوزن الولادي لجميع أطفال المجتمع حديثي الولادة ، وبالتالي فإنه يمكننا القول أن هذا التوزع الولادي سوف يخضع تماما لشكل التوزع الطبيعي .

- وإن لهذا التوزع الطبيعي مجموعة من الخصائص التي سنذكرها بعد قليل ...

التوزع الطبيعي

- Total area under the curve = 1 (or 100%).
- Bell shaped and symmetrical about its mean.
- The peak of the curve lies above the mean.
- Any position along the horizontal axis can be expressed as a number of SDs away from the mean.
- The mean and median coincide.

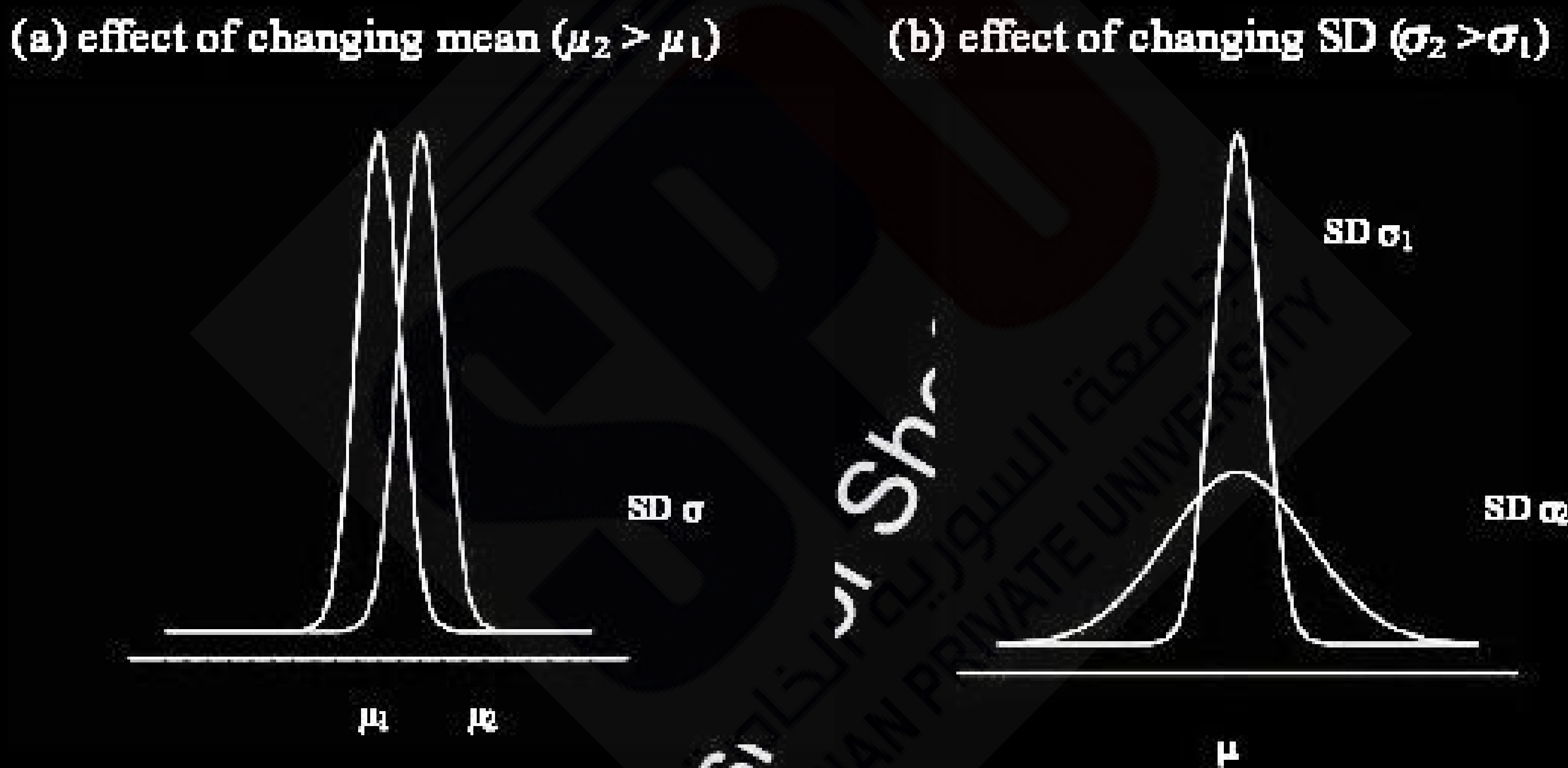


التوزع الطبيعي

- يتم وصف التوزع الطبيعي بشكل كامل من خلال المعلمتين التاليتين : أولا : μ والذي يعبر عن الوسط الحسابي للمجتمع الأم أو مركزية التوزيع ، ثانيا : σ الانحراف المعياري للمجتمع أو المعبر عن مقدار التشتت في المجتمع المدروس.

- عندما يتسم المجتمع المدروس بأن قيمة σ صغيرة فإن التوزع يتسم بقربه من مركزه المسمى μ ، ولكن عندما تكون σ كبيرة فإن مشاهدات التوزع تكون منتشرة على طول محور القياس

التوزع الطبيعي



التوزع الطبيعي

- هناك عدد غير محدود من التوزعات الطبيعية اعتمادا على قيمة كل من μ ، σ . ولكن «التوزع الطبيعي المعياري» Standard Normal Distribution هو حالة خاصة من حالات التوزع الطبيعي يكون فيها الوسط الحسابي للمجتمع الأم = صفر ، الانحراف المعياري هو ١ (أو التباين = ١).

- إذا كان المتغير العشوائي X يخضع للتوزع الطبيعي ، مع وسط حسابي μ ، وبانحراف معياري σ فإن المشتق z الطبيعي المعياري (وهو يساوي المتغير X - μ مقسوما على σ) هو متغير عشوائي يخضع للتوزع الطبيعي المعياري .

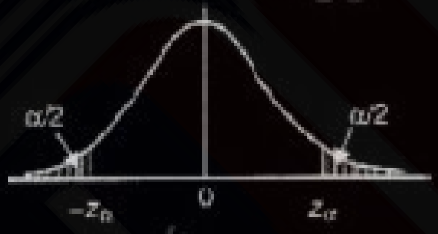
التوزع الطبيعي

- إن المساحات المحصورة تحت منحنى التوزع الطبيعي المعياري قد تمت جدولتها بشكل كامل ...

Table 5.1 Selected probabilities associated with the Normal distribution

Standardised deviate	Probability of greater deviation	
$z = (x - \mu)/\sigma$	In either direction	In one direction
0	1.000	0.500
1	0.317	0.159
2	0.046	0.023
3	0.0027	0.0013
1.645	0.10	0.05
1.960	0.05	0.025
2.576	0.01	0.005

Table T1 The Normal distribution. The value tabulated is the probability, x, that a random variable. Normally distributed with mean zero and standard deviation one, will be greater than z or less than $-z$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	1.0000	0.9920	0.9840	0.9761	0.9681	0.9601	0.9522	0.9442	0.9362	0.9283
0.10	0.9203	0.9124	0.9045	0.8966	0.8887	0.8808	0.8729	0.8650	0.8572	0.8493
0.20	0.8415	0.8337	0.8259	0.8181	0.8103	0.8206	0.7949	0.7872	0.7795	0.7718
0.30	0.7642	0.7566	0.7490	0.7414	0.7339	0.7263	0.7188	0.7114	0.7039	0.6965
0.40	0.6892	0.6818	0.6745	0.6672	0.6599	0.6527	0.6455	0.6384	0.6312	0.6241
0.50	0.6171	0.6101	0.6031	0.5961	0.5892	0.5823	0.5755	0.5687	0.5619	0.5552
0.60	0.5485	0.5419	0.5353	0.5287	0.5222	0.5157	0.5093	0.5029	0.4965	0.4902
0.70	0.4839	0.4777	0.4715	0.4654	0.4593	0.4533	0.4473	0.4413	0.4354	0.4295
0.80	0.4237	0.4179	0.4122	0.4065	0.4009	0.3953	0.3898	0.3843	0.3789	0.3735
0.90	0.3681	0.3628	0.3576	0.3524	0.3472	0.3421	0.3371	0.3320	0.3271	0.3222
1.00	0.3173	0.3125	0.3077	0.3030	0.2983	0.2837	0.2891	0.2846	0.2801	0.2757

TABLE T1

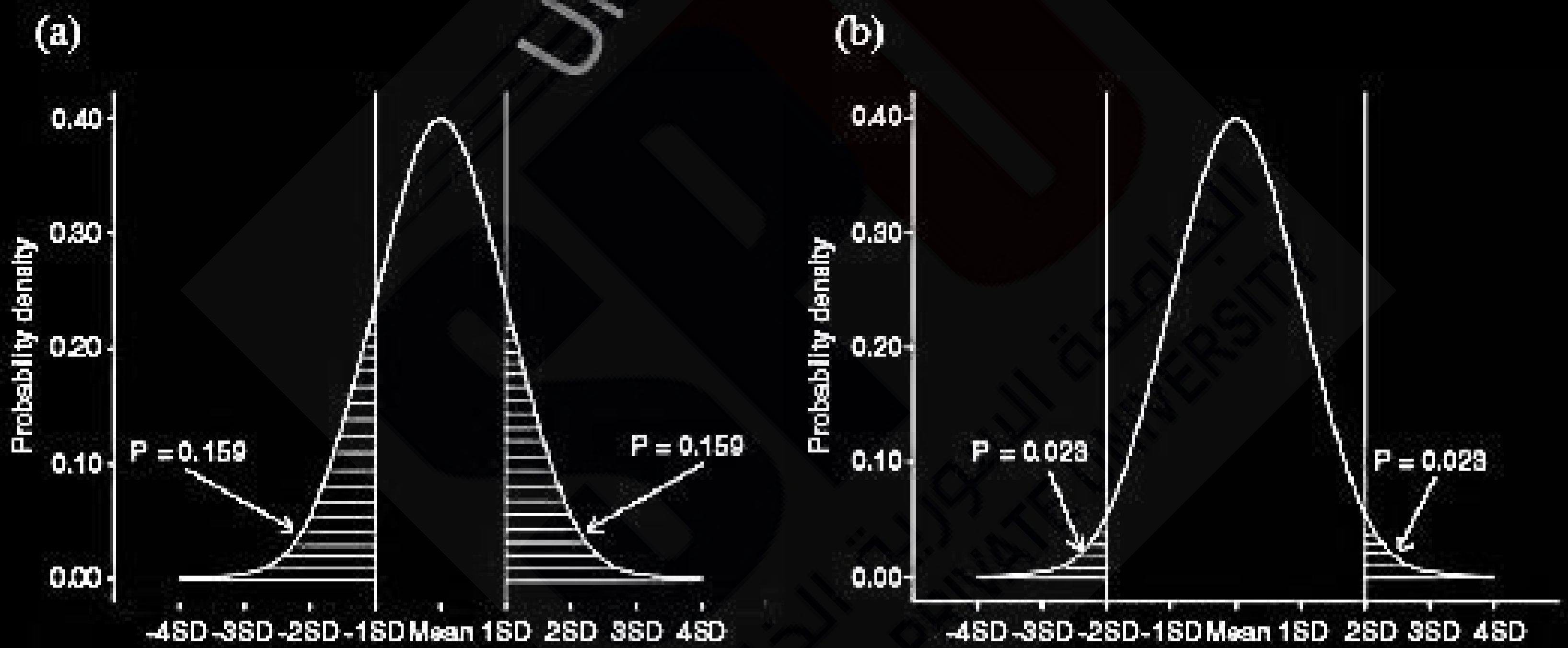
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.00	0.3173	0.3125	0.3077	0.3030	0.2983	0.2937	0.2891	0.2846	0.2801	0.2757
1.10	0.2713	0.2670	0.2627	0.2585	0.2543	0.2501	0.2460	0.2420	0.2380	0.2340
1.20	0.2301	0.2263	0.2225	0.2187	0.2150	0.2113	0.2077	0.2041	0.2005	0.1971
1.30	0.1936	0.1902	0.1868	0.1835	0.1802	0.1770	0.1738	0.1707	0.1676	0.1645
1.40	0.1615	0.1585	0.1556	0.1527	0.1499	0.1471	0.1443	0.1416	0.1389	0.1362
1.50	0.1336	0.1310	0.1285	0.1260	0.1236	0.1211	0.1188	0.1164	0.1141	0.1118
1.60	0.1096	0.1074	0.1052	0.1031	0.1010	0.0989	0.0969	0.0949	0.0930	0.0910
1.70	0.0891	0.0873	0.0854	0.0836	0.0819	0.0801	0.0784	0.0767	0.0751	0.0735
1.80	0.0719	0.0703	0.0688	0.0672	0.0658	0.0643	0.0629	0.0615	0.0601	0.0588
1.90	0.0574	0.0561	0.0549	0.0536	0.0524	0.0512	0.0500	0.0488	0.0477	0.0466
2.00	0.0455	0.0444	0.0434	0.0424	0.0414	0.0404	0.0394	0.0385	0.0375	0.0366

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.00	0.0455	0.0444	0.0434	0.0424	0.0414	0.0404	0.0394	0.0385	0.0375	0.0366
2.10	0.0357	0.0349	0.0340	0.0332	0.0324	0.0316	0.0308	0.0300	0.0293	0.0285
2.20	0.0278	0.0271	0.0264	0.0257	0.0251	0.0244	0.0238	0.0232	0.0226	0.0220
2.30	0.0214	0.0209	0.0203	0.0198	0.0193	0.0188	0.0183	0.0178	0.0173	0.0168
2.40	0.0164	0.0160	0.0155	0.0151	0.0147	0.0143	0.0139	0.0135	0.0131	0.0128
2.50	0.0124	0.0121	0.0117	0.0114	0.0111	0.0108	0.0105	0.0102	0.0099	0.0096
2.60	0.0093	0.0091	0.0088	0.0085	0.0083	0.0080	0.0078	0.0076	0.0074	0.0071
2.70	0.0069	0.0067	0.0065	0.0063	0.0061	0.0060	0.0058	0.0056	0.0054	0.0053
2.80	0.0051	0.0050	0.0048	0.0047	0.0045	0.0044	0.0042	0.0041	0.0040	0.0039
2.90	0.0037	0.0036	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028
3.00	0.0027	0.0026	0.0025	0.0024	0.0024	0.0023	0.0022	0.0021	0.0021	0.0020

التوزع الطبيعي

- من الجدول البسيط السابق ، نلاحظ أنه من أجل قيمة z (والتي هي عبارة عن عدد مرات الانحراف المعياري بعدا عن الوسط الحسابي المساوي للصفر) نستطيع حساب المساحة تحت المنحنى المتبقية والتي تقع إلى اليمين أو إلى اليسار من هذه القيمة (أي باتجاه واحد) أو حتى المساحة المشتركة (أي بكلا الاتجاهين) ... .

التوزع الطبيعي

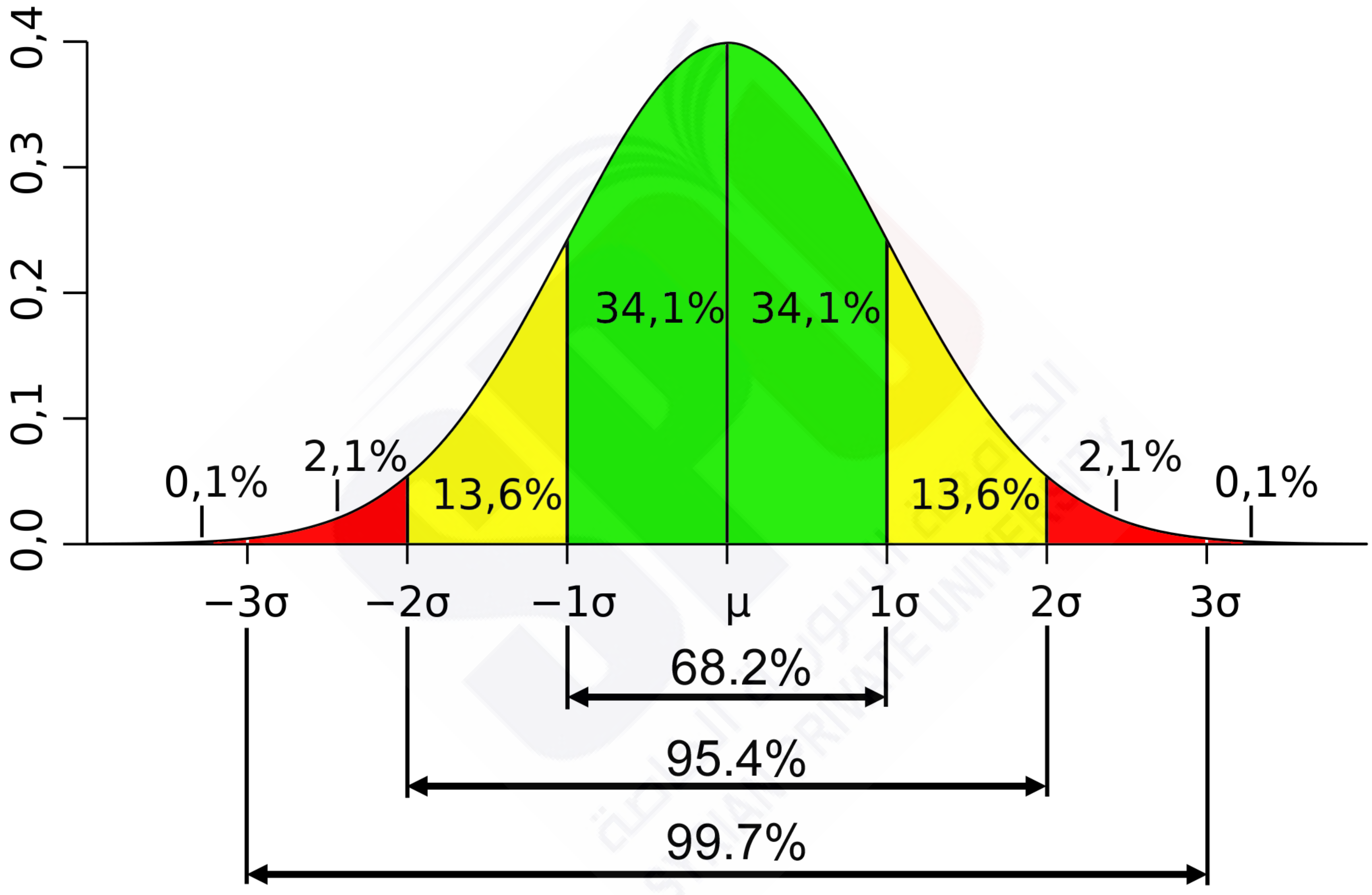


التوزع الطبيعي

- من خلال الشكل السابق نستطيع ملاحظة أنه ٦٨% من الاحتمالية تقع بين +١ و -١ انحراف معياري SD ،

- وأن الأغلبية العظمي (أي ٩٥%) من الحالات يمكن أن تقع بين +٢ ، -٢ انحراف معياري SD ،

- وتقريبا كل (أي حوالي ٩٩%) من المشاهدات تقع بين +٣ و -٣ انحراف معياري SD عن الوسط الحسابي .


0,4
0,3
0,2
0,1
0,0
34,1%
34,1%
13,6%
13,6%
2,1%
2,1%
0,1%
0,1%
−3σ
−2σ
−1σ
μ
1σ
2σ
3σ
68.2%
95.4%
99.7%

التوزع الطبيعي

- من خلال قيمة z نستطيع حساب المساحات المحصورة تحت منحنى التوزع الطبيعي ..

As can be seen from Table 5.1 using, z values of 1.96, that is, 1.96 SD away from the mean) then exactly 95% of the Normal distribution lies between

$$\mu - 1.96 \times \sigma \text{ and } \mu + 1.96 \times \sigma.$$

- وبرفع قيمة z إلى ٢.٥٧٦ فإن ٩٩% من التوزع الطبيعي يقع في المجال الموافق . في الممارسة العملية فإن المعلمتين الإحصائيتين للمجتمع وهما **μ** ، **σ** ، يتم تخمينهما من خلال بيانات العينة Sample المسحوبة من المجتمع . وبالتالي أول خطوة يقوم بها الباحث هو سحب عينة عشوائية من المجتمع المدروس ...

التوزع الطبيعي

كيف لنا أن نستخدم التوزع الطبيعي ؟

- إن التوزع الاحتمالي الطبيعي يمكن أن يستخدم من أجل حساب احتمالية حدوث قيم مختلفة ...
- قد نكون مهتمين بالسؤال التالي : ما هو احتمالية أن نقع ضمن مجال انحراف معياري واحد عن الوسط الحسابي أو خارج هذا النطاق ؟

- يمكن لنا الاستفادة من جدول التوزع الطبيعي لكي نعلم احتمالية وقوعنا خارج النطاق المحدد ..

التوزع الطبيعي

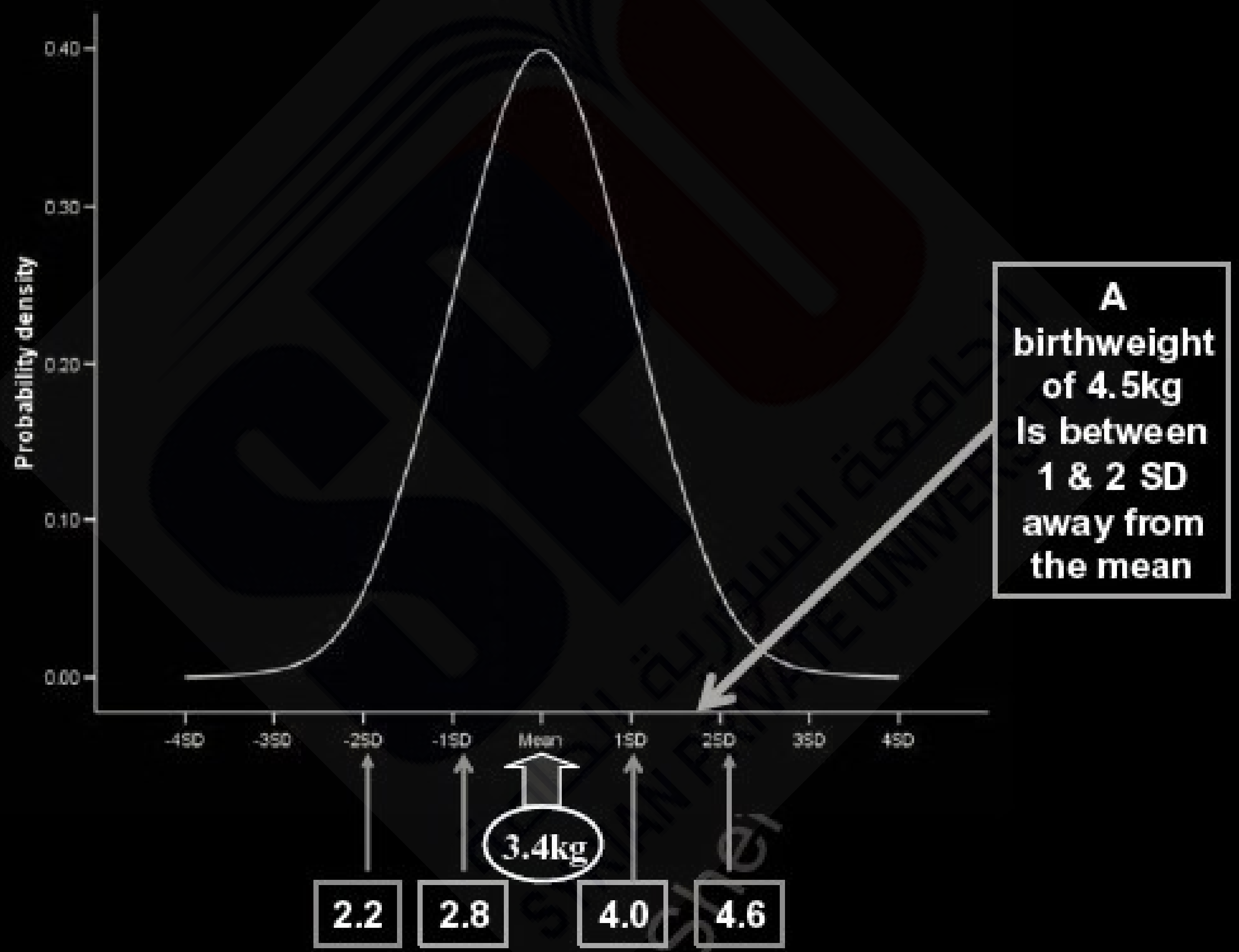
- باستخدام الدراسة المذكورة سابقا O'Cathain وآخرين لعام ٢٠٠٢ ، حول الوزن الولادي ولنفترض أن هذا المتغير يخضع للتوزع الطبيعي ، بوسط حسابي ٣.٤ كغ و انحراف معياري ٠.٦ كغ ، السؤال : ما هو احتمالية أن يلد طفل بوزن ٤.٥ كغ أو أكثر ؟

التوزع الطبيعي

Since birthweight is assumed to follow a Normal distribution, with mean of 3.4kg and SD of 0.6kg, we therefore know that approximately 68% of birthweights will lie between 2.8 and 4.0kg and about 95% of birthweights will lie between 2.2 and 4.6kg. Using Figure 5.9 we can see that a birthweight of 4.5kg is between 1 and 2 standard deviations away from the mean.

First calculate, z, the number of standard deviations 4.5kg is away from the mean of 3.4kg, that is, $z = \frac{4.5 - 3.4}{0.6} = 1.83$. Then look for $z = 1.83$ in Table T1 of the Normal distribution table which gives the probability of being outside the values of the mean −1.83 SD to mean +1.83 SD as 0.0672. Therefore the probability of having a birthweight of 4.5kg or higher is 0.0672/2 = 0.0336 or 3.3%.

التوزع الطبيعي



التوزع الطبيعي

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.00	0.3173	0.3125	0.3077	0.3030	0.2983	0.2937	0.2891	0.2846	0.2801	0.2757
1.10	0.2713	0.2670	0.2627	0.2585	0.2543	0.2501	0.2460	0.2420	0.2380	0.2340
1.20	0.2301	0.2263	0.2225	0.2187	0.2150	0.2113	0.2077	0.2041	0.2005	0.1971
1.30	0.1936	0.1902	0.1868	0.1835	0.1802	0.1770	0.1738	0.1707	0.1676	0.1645
1.40	0.1615	0.1585	0.1556	0.1527	0.1499	0.1471	0.1443	0.1416	0.1389	0.1362
1.50	0.1336	0.1310	0.1285	0.1260	0.1236	0.1211	0.1188	0.1164	0.1141	0.1118
1.60	0.1096	0.1074	0.1052	0.1031	0.1010	0.0989	0.0969	0.0949	0.0930	0.0910
1.70	0.0891	0.0873	0.0854	0.0836	0.0819	0.0801	0.0784	0.0767	0.0751	0.0735
1.80	0.0719	0.0703	0.0688	$\boxed{0.0672}$	0.0658	0.0643	0.0629	0.0615	0.0601	0.0588
1.90	0.0574	0.0561	0.0549	0.0536	0.0524	0.0512	0.0500	0.0488	0.0477	0.0466
2.00	0.0455	0.0444	0.0434	0.0424	0.0414	0.0404	0.0394	0.0385	0.0375	0.0366

TABLE T1

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.00	0.0455	0.0444	0.0434	0.0424	0.0414	0.0404	0.0394	0.0385	0.0375	0.0366
2.10	0.0357	0.0349	0.0340	0.0332	0.0324	0.0316	0.0308	0.0300	0.0293	0.0285
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2.90	0.0037	0.0036	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028
3.00	0.0027	0.0026	0.0025	0.0024	0.0024	0.0023	0.0022	0.0021	0.0021	0.0020

Table T1 of the Normal distribution table which gives the probability of being outside the values of the mean –1.83 SD to mean +1.83 SD as 0.0672. Therefore the probability of having a birthweight of 4.5 kg or higher is 0.0672/2 = 0.0336 or 3.3%.

التوزع الطبيعي

- هناك استخدامات أخرى للتوزع الطبيعي في الإحصاء الحيوي ، وهو يستخدم كمقاربة للتوزع ثنائي الحد (الحداني) ، و توزع بواسون ...

- إن **التوزع ثنائي الحد** لأي باي π يمكن أن يقترب من التوزع الطبيعي بزيادة n (عدد الأفراد). والوصول للتوزع الطبيعي أكثر سرعة عندما تكون باي قريبة من ٠.٥ ، وعندما تكون عدد المشاهدات كبيرة (أي n كبيرة) يمكن لمتغير ثنائي أن يعتبر خاضعا للتوزع الطبيعي بوسط حسابي هو n*π ، وانحراف معياري هو الجذر التربيعي لحاصل ضرب nπ في (1-π) .

- وبالنسبة **لتوزع بواسون** مع وسط λ ، فإنه يقترب من التوزع الطبيعي مع زيادة قيمة λ . وعندما تكون لامدا كبيرة يمكن لتوزع بواسون أن يعتبر «توزعا طبيعيا» بوسط حسابي لامدا ، وانحراف معياري هو الجذر التربيعي لــ λ .

شكرا جزيلا لحسن استماعكم

myhajeer@gmail.com