Quantitative Methods – I

Practice 10

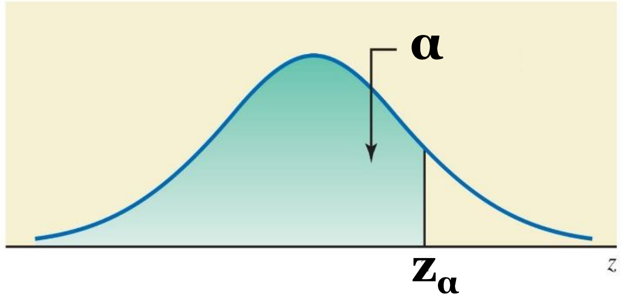
Lorenzo Cavallo

For any clarification:
lorenzo.cavallo.480084@uniroma2.eu

THEME #1

How to use the Table of Standard Normal distribution

How to use the Table of Standard Normal distribution



z_α =-1.65

α =0.0495=4.95%~5%

Area [-∞;-1.65]=0.0495
P(x<z_α) =0.0495=4.95%

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

z_α
Standardized value

α
Shaded Area [-∞;z_α]
Probability P(x)< z_α

α =0.025=2.5%

Find the nearest value to 0.025 in the Table

The value of z for the α is -1.96

$z_\alpha = z_{0.025}$ =-1.96

Area [-∞;-1.96]=0.025
the 2.5% of the values are lower than -1.96

α=0.01 → z_{α}=-2.33

α=0.025 → z_{α}=-1.96

α=0.05 → z_{α}=-1.64 or -1.65

α=0.1 → z_{α}=-1.28

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379

$\alpha=0.99 \rightarrow z_\alpha=2.33$

$\alpha=0.975 \rightarrow z_\alpha=1.96$

$\alpha=0.95 \rightarrow z_\alpha=1.64$ or 1.65

$\alpha=0.90 \rightarrow z_\alpha=1.28$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916

THEME #2

—

Confidence Intervals

Inferential Statistics

"making decision on a population based on sample results"

Two main tools

Estimation

Hypothesis Testing

Point Estimate

Estimate a **parameter**:
Ex: μ

by an **estimator**:
Ex: $\bar{x}$

Interval of Confidence

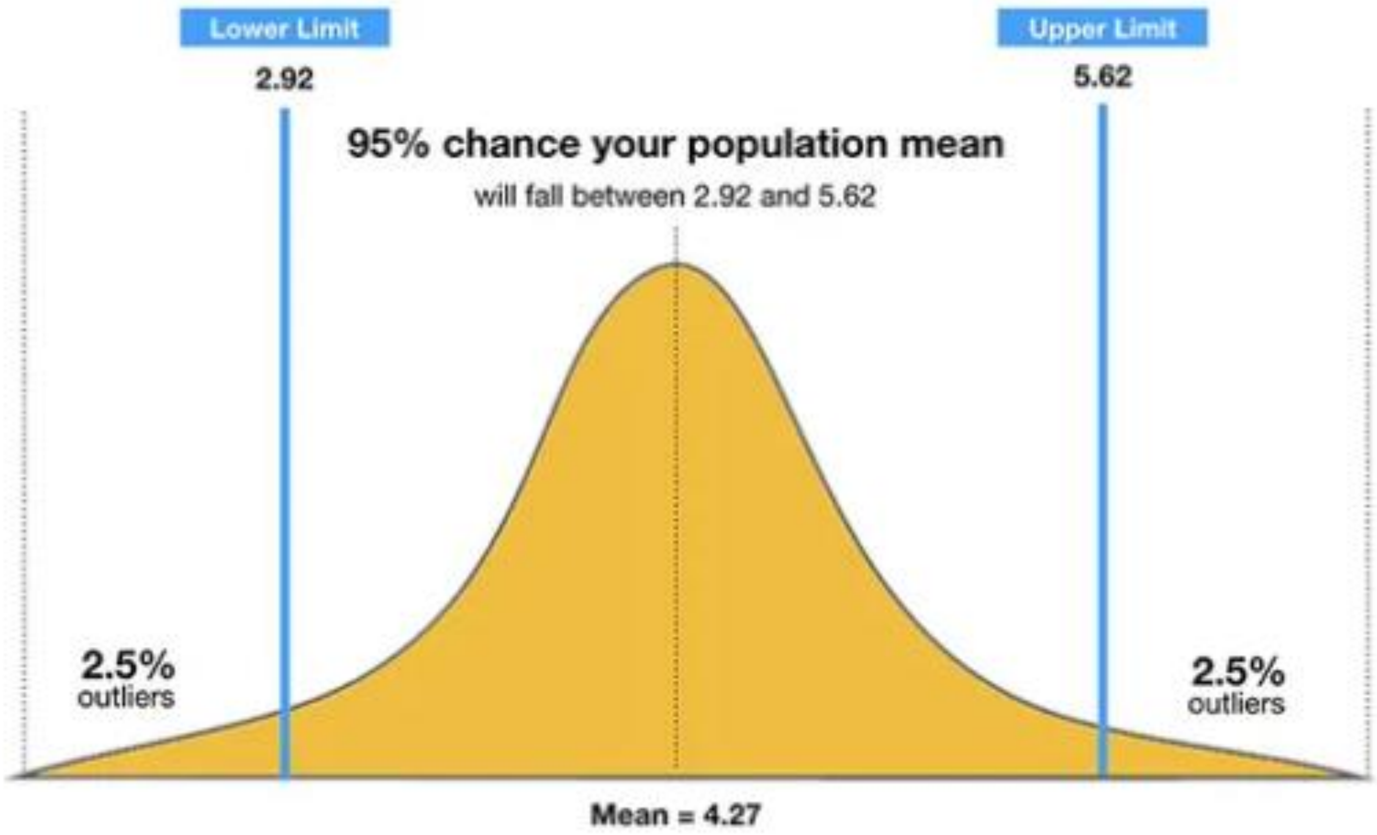
Estimate a **parameter**:
Ex: μ

by an **interval** in which we have **confidence** that the parameter fall:
Ex: $\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$

Find values (Estimation) or Test statements (Hypothesis Tests) on population parameters based on sample estimator

What are Confidence Intervals in Statistics?

The confidence interval (CI) is a range of values that's likely to include a population value with a certain degree of confidence. It is often expressed a % whereby a population mean lies between an upper and lower interval.



Confidence Interval

The confidence interval is given as

$$\text{Point estimate} \pm \text{Margin of error}$$

Ex. the (1-α)% confidence interval for the population mean μ is:

$\bar{x}$ is the **point estimate** of the population mean μ

$$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$$

is called **margin of error**

$\bar{x}$ is the sample mean (the point estimate of the population mean)

$\sigma_{\bar{x}}$ is the standard error $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ (the standard deviation of the sample mean)

$z_{\alpha/2}$ is a z-score

The **confidence level (1-α)** associated with a confidence interval states how much confidence we have that this interval contains the true population parameter.

The confidence level is denoted by (1-α)% and express a probability.

Confidence Level and Confidence Interval

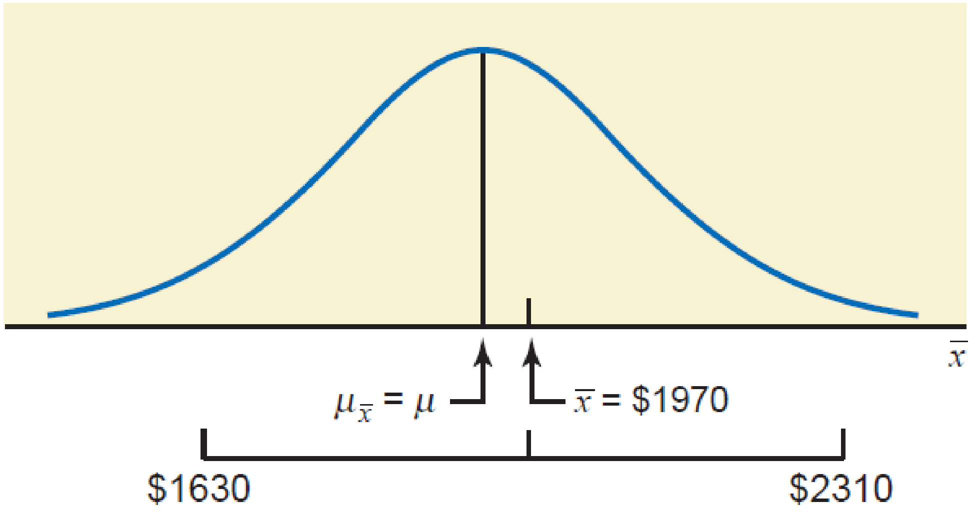
The confidence interval is Point estimate $\pm$ Margin of error

The (1-α)% confidence interval for the population mean μ is:

$$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$$

Ex.

We have a sample mean of $1970



We have a statistical confidence of the 95% that the population mean μ lies in the interval from $ 1630 and $2310.

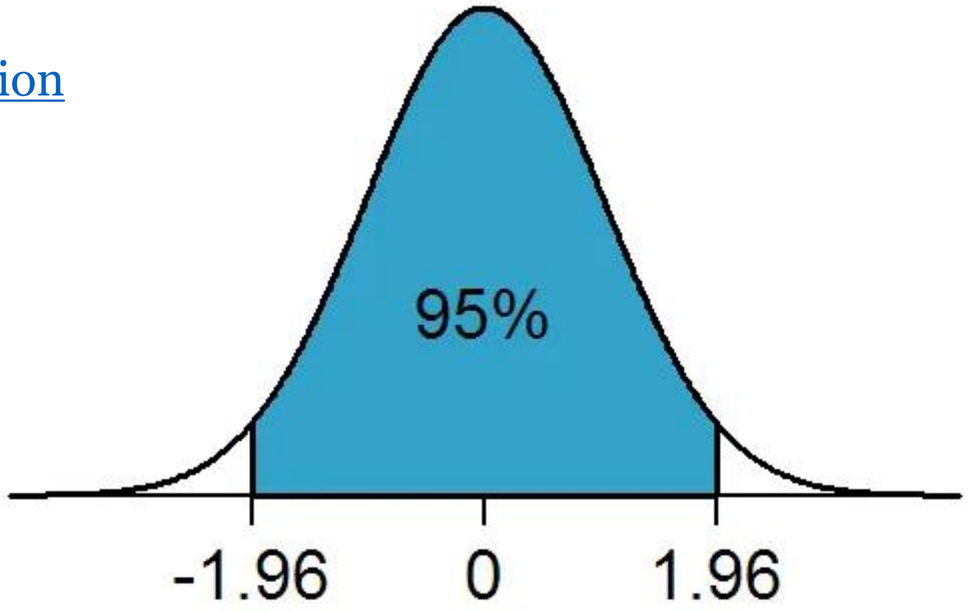
In other terms we have a 95% probability that population mean lies in the interval: [$ 1630 ; $2310].

What does a 95% confidence interval mean?

The 95% confidence interval is a range of values that you can be 95% certain contains the true mean of the population.

As the sample size increases, the range of interval values will narrow, meaning that you know that mean with much more accuracy compared with a smaller sample.

We can visualize this using a standard normal distribution



For example, the probability of the population mean value being between -1.96 and +1.96 standard deviations (z-scores) from the sample mean is 95%.

Accordingly, there is a 5% chance that the population mean lies outside of the upper and lower confidence interval (as illustrated by the 2.5% of outliers on either side of the 1.96 z-scores).

If the standard deviation of the population (σ) is known the standard error (the standard deviation of the sample mean) **is:**

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

If the standard deviation of the population (σ) is NOT known the standard error (the standard deviation of the sample mean) **is:**

$$\sigma_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}}$$

$\hat{\sigma}$ (or sometimes is noted with the letter «s») is the sample standard deviation and the unbiased estimator of the standard deviation

If the standard deviation of the population (σ) is known the margin of error is

$$\boldsymbol{z_{\alpha/2}\sigma_{\bar{x}}} = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

where $z_{\alpha/2}$ is a z score
and $\frac{\sigma}{\sqrt{n}}$ is the standard error (with σ known)

If the standard deviation of the population (σ) is NOT known the margin of error is

$$\boldsymbol{t_{n-1;\alpha/2}\sigma_{\bar{x}}} = t_{n-1;\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$$

where $t_{n-1;\alpha/2}$ is a Student's T with n-1 degrees of freedom
and $\hat{\sigma}$ is the sample standard deviation (with σ not known)

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)$$

X Normal r.v. ~$N(\mu\,;\sigma^2)$?

- Yes → σ known?
  - Yes →
    $$(1-\alpha)\%IC(\mu)=\left[\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\,;\bar{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$
  - No →
    $$(1-\alpha)\%IC(\mu)=\left[\bar{x}-t_{n-1;\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}\,;\bar{x}+t_{n-1;\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}\right]$$
- No → n ≥ 25?
  - No → NO Interval of Confidence
  - Yes, apply the **Central Limit Theorem** → σ known?
    - Yes →
      $$(1-\alpha)\%IC(\mu)\cong\left[\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\,;\bar{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$
    - No →
      $$(1-\alpha)\%IC(\mu)\cong\left[\bar{x}-z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}\,;\bar{x}+z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}\right]$$

X Normal r.v.	
σ known	$$(1-\alpha)\%IC(\mu) = \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}};\bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$
σ unknown	$$(1-\alpha)\%IC(\mu) = \left[\bar{x} - t_{n-1;\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}};\bar{x} + t_{n-1;\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}\right]$$

X not Normal r.v. (apply the Central Limit Theorem)	
σ known	$$(1-\alpha)\%IC(\mu) \cong \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}};\bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$
σ unknown	$$(1-\alpha)\%IC(\mu) \cong \left[\bar{x} - z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}};\bar{x} + z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}\right]$$

X Bernoullian r.v., π unknown	
σ known	$$(1-\alpha)\%IC(\pi) \cong \left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}};\hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}\right]$$

Confidence Level	Z-Score
0.90	1.645
0.95	1.96
0.99	2.58

Exercise 1b

A company has made a new smartphone. Before the company decides the price at which to sell the smartphone, it wants to know the average price of all smartphones in the market. The research department took a sample of **50** comparable smartphones and collected information on their prices. This information produced a **mean price** of $**145** for this sample. It is known that the **standard deviation** of the prices of all smartphones is $**35** and the population of such prices is **normal**.

a. What is the point estimate of the mean price of all smartphones?
b. Calculate the standard deviation of the sample mean.
c. Construct a 90% confidence interval for the mean price of all smartphones. Which is the margin of error?
d. Construct a 95% confidence interval for the mean price of all smartphones.
e. The research department add to the first sample other 50 smartphones with a mean price of $165. Calculate the new point estimate of the mean price and the 95% confidence interval for the mean.

Solution

From the given information, n=50 $\bar{x}$=145 and standard deviation is 35 (σ known).

a. The point estimate of the mean price is $145 (the sample mean $\bar{x}$)

b. The standard deviation of the sample mean (standard error) with σ known is:

$$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 35/\sqrt{50} = 35 \,/\, 7.07 = 4.95$$

Exercise 1b

A company has made a new smartphone. Before the company decides the price at which to sell the smartphone, it wants to know the average price of all smartphones in the market. The research department took a sample of 50 comparable smartphones and collected information on their prices. This information produced a mean price of \$145 for this sample. It is known that the standard deviation of the prices of all smartphones is \$35 and the population of such prices is normal.

c. Construct a 90% confidence interval for the mean price of all smartphones. Which is the margin of error?

Solution

c. The confidence level is 90%.

The z-value (from the Table of the standard normal distribution) for a 90% confidence level is 1.65

The 90% confidence interval for the mean μ is:

$$(1-\alpha)\%IC(\mu) = \left[\bar{x} - z\alpha/_2 \frac{\sigma}{\sqrt{n}} ; \bar{x} + z\alpha/_2 \frac{\sigma}{\sqrt{n}}\right]$$

4.95

$$90\%IC(\mu) = \left[145 - 1.65\frac{35}{\sqrt{50}} ; 145 + 1.65\frac{35}{\sqrt{50}}\right] = [136.83\ ;\ 153.17]$$

We are 90% confident that the mean price of all smartphones is between \$136.83 and \$153.17.

The margin of error is $z\sigma_{\bar{x}} = 1.65 \cdot 4.95 = 8.1675$

Exercise 1b

A company has made a new smartphone. Before the company decides the price at which to sell the smartphone, it wants to know the average price of all smartphones in the market. The research department took a sample of 50 comparable smartphones and collected information on their prices. This information produced a mean price of $145 for this sample. It is known that the standard deviation of the prices of all smartphones is $35 and the population of such prices is normal.

d. Construct a 95% confidence interval for the mean price of all smartphones.

Solution

d. The confidence level is 95%. The z value for a 95% confidence level is 1.96

The 95% confidence interval for the mean μ is:

$$(1-\alpha)\%IC(\mu) = \left[\bar{x} - z\alpha/_2 \frac{\sigma}{\sqrt{n}} ; \bar{x} + z\alpha/_2 \frac{\sigma}{\sqrt{n}}\right]$$

$$95\%IC(\mu) = \left[145 - 1.96\frac{35}{\sqrt{50}} ; 145 + 1.96\frac{35}{\sqrt{50}}\right] = [135.298 ; 154.702]$$

We are 95% confident that the mean price of all smartphones is between $135.298 and $154.702.

Exercise 1b

A company has made a new smartphone. Before the company decides the price at which to sell the smartphone, it wants to know the average price of all smartphones in the market. The research department took a sample of 50 comparable smartphones and collected information on their prices. This information produced a mean price of $145 for this sample. It is known that the standard deviation of the prices of all smartphones is $35 and the population of such prices is normal.

e. The research department add to the first sample **other 50** smartphones with a mean price of $**165**. Calculate the new point estimate of the mean price and the 95% confidence interval for the mean.

Solution

e. The mean of the first sample is $\bar{x}_1$=145 and the mean of the second sample is $\bar{x}_2$=165

$$\bar{x}_{1+2} = \frac{\bar{x}_1 \cdot n_1 + \bar{x}_2 \cdot n_2}{100} = \frac{50 \cdot 145 + 50 \cdot 165}{100} = \mathbf{155}$$

The standard error is $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 35/\sqrt{100} = 35 / 10 = \mathbf{3.5}$

The confidence level is 95%. The z value for a 95% confidence level is 1.96

The 95% confidence interval for the mean is

$$(1-\alpha)\% IC(\mu) = \left[\bar{x} - z\alpha/_2 \frac{\sigma}{\sqrt{n}} ; \bar{x} + z\alpha/_2 \frac{\sigma}{\sqrt{n}}\right]$$

$$95\% IC(\mu) = [155 - 1.96 \cdot 3.5 ; 155 + 1.96 \cdot 3.5] = [148.14 ; 161.86]$$

We are 95% confident that the mean price of all smartphones is between $148.14 and $161.86.

Exercise 1

For a data set obtained from a sample, **n = 20** and **$\bar{x}$ = 35**. It is known that the **population standard deviation** is **3.1** and the population is **normally** distributed.

a. What is the point estimate of μ?

b. Make a 99% confidence interval for μ.

c. What is the margin of error?

d. What is the margin of error of the 95% confidence interval?

Solution

We know that the population is normally distributed, so

$$X \sim N(\mu\ ;\sigma^2)$$

with the unknow mean μ and a standard deviation σ = 3.1

$$X \sim N(\mu\ ;\sigma^2 = 3.1^2 = 9.61)$$

Also the sample mean is normally distributed,

$$\bar{x} \sim N(\mu\ ;\frac{\sigma^2}{n})$$

$$\bar{x} \sim N(\mu\ ;\frac{\sigma^2}{n} = \frac{3.1^2}{20} = 0.4805)$$

If $\frac{\sigma^2}{n} = 0.4805$ the standard error $\frac{\sigma}{\sqrt{n}} = \frac{3.1}{\sqrt{20}} = 0.6932$

Exercise 1

For a data set obtained from a sample, n = 20 and $\bar{x}$ = 35. It is known that the population standard deviation is 3.1 and the population is normally distributed.

a. What is the point estimate of μ?

b. Make a 99% confidence interval for μ.

c. What is the margin of error?

d. What is the margin of error of the 95% confidence interval?

Solution

a. The point estimate of the population mean (μ) is the sample mean $\bar{x}$ = 35

b. We have to calculate the $(1-\alpha)\%IC(\mu)$ for the population mean, μ, for the confidence levels of $(1-\alpha) = 0.99$

If (1- α)=0.99, α=0.01 and $^{\alpha}/_{2}$=0.005

The $z_{\alpha/2}$ (in absolute value) from the table of the standard normal distribution for $^{\alpha}/_{2}$=0.005 is

$$\left|z_{\alpha/2}\right| = |z_{0.005}| = 2.58$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049

A little tip

To find the value of the $z_{\alpha/2}$ you use the table of the standard normal distribution.

If you have the negative table of Z the $z_{\alpha/2}$ will be (in general) negative.

You have to use it in absolute value.

Ex. For (1- α)=0.99, α=0.01 and $\alpha/_2$=0.005

$$\left|z_{\alpha/2}\right| = |z_{0.005}| = 2.58$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049

If you have only the positive table of Z you have α values only ≥ 0.5.

To find the correct Z-score you have to find the symmetric and complementary value of $\alpha/_2$: (1- $\alpha/_2$)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952

Ex. For (1- α)=0.99, α=0.01 and $\alpha/_2$=0.005

$1 - \alpha/_2$=0.995

$$z_{1-\alpha/2} = z_{0,995} = 2.58$$

Exercise 1

For a data set obtained from a sample, n = 20 and $\bar{x}$ = 35. It is known that the population standard deviation is 3.1 and the population is normally distributed.

a. What is the point estimate of μ?

b. Make a **99**% confidence interval for μ.

c. What is the margin of error?

d. What is the margin of error of the **95**% confidence interval?

Solution

The IC formula for X normally distributed and σ known is:

$$(1-\alpha)\%IC(\mu) = \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

We have: $\bar{x}$= 35 , n=20 , σ = 3.1

The standard error (or the standard deviation of the sample mean) is: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ = 3.1/√20 = 0.69

99%IC(μ)=[$\bar{x}$ - $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$; $\bar{x}$ + $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$] = [35 – 2.58 · 0.69 ; 35 + 2.58 · 0.69] = [33.23 ; 36.77]

c. The margin of error is $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ = 2.58 · 0.69 = 1.77

d. The margin of error is $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$. The difference with c. is only for the z score $z_{\alpha/2}$ that for $(1-\alpha) =$ 0.95 is $z_{\alpha/2} = z_{0.025}$ = 1.96

The margin of error is $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ = 1.96 · 0.69 = 1.35

Exercise 2

A machine is set up such that the average content of juice per bottle is equals to μ with a **variance** of **25** cl.

A sample of **100** bottles yields an average content of **48**cl.

Calculate a 90% and a 95% confidence interval for the average content.

Solution

The average content of juice per bottle of the machine is the variable X with the unknown mean μ and a variance $\sigma^2 = 25$. The standard deviation of the machine is the square root of the variance, so $\sigma = 5$

It is **not known** if the **X** is **normally distributed**, **but** the sample of bottles is large enough (**n=100>25**) to apply the Central Limit Theorem.

For the Central Limit Theorem,

$$\bar{x} \sim N(\mu \,; \frac{\sigma^2}{n})$$

We have to calculate the $(1-\alpha)\%IC(\mu)$ for the population mean, μ, for the confidence levels of $(1-\alpha) = 0.90$ and $(1-\alpha) = 0.95$

For (1- α)=0.90, α=0.1 and $^{\alpha}/_{2}$=0.05. The $z_{\alpha/2}$ for $(1-\alpha) = 0.90$ is $z_{\alpha/2} = z_{0.05} = 1.645$

For (1- α)=0.95, α=0.05 and $^{\alpha}/_{2}$=0.025. The $z_{\alpha/2}$ for $(1-\alpha) = 0.95$ is $z_{\alpha/2} = z_{0.025} = 1.96$

Exercise 2

A machine is set up such that the average content of juice per bottle is equals to μ with a variance of 25 cl.
A sample of 100 bottles yields an average content of 48cl.
Calculate a 90% and a 95% confidence interval for the average content.

Solution

The IC formula for X not normally distributed and σ known is:

$$(1-\alpha)\%IC(\mu) \cong \left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} ; \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

We have: $\bar{x}$= 48 , n=100 , σ = 5

The standard error is: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ = 5/√100 = 5 / 10 = 0.50

The $z_{\alpha/2}$ for $(1-\alpha) = 0.90$ is $z_{\alpha/2} = z_{0.05}$= 1.645

The $z_{\alpha/2}$ for $(1-\alpha) = 0.95$ is $z_{\alpha/2} = z_{0.025}$= 1.96

90%IC(μ) ≅ [$\bar{x}$ - 1.64 $\frac{\sigma}{\sqrt{n}}$; $\bar{x}$ + 1.64 $\frac{\sigma}{\sqrt{n}}$] = [48 – 1.645 · 0.5 ; 48 + 1.645 · 0.5] = [47.18 ; 48.82]

95%%IC(μ) ≅ [$\bar{x}$ - 1.96 $\frac{\sigma}{\sqrt{n}}$; $\bar{x}$ + 1.96 $\frac{\sigma}{\sqrt{n}}$] = [48 - 1.96 · 0.5 ; 48 + 1.96 · 0.5] = [47.02 ; 48.98]

Confidence Interval for μ (σ not known)

The (1-α)% ***confidence interval for μ*** is:

$$\bar{x} \pm t\sigma_{\bar{x}}$$

where

$$\sigma_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}} \text{ (the standard error)}$$

with

$$\sigma_{\bar{x}} = s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}} \quad \text{the standard deviation for a sample}$$

The value of *t* used here is obtained from the *t* distribution table for *n-1* degrees of freedom and the given confidence level.

$E = t\sigma_{\bar{x}}$ is the **margin of error**

Exercise 3

A sample of **30** products selected from a **normally** distributed population has a **mean** weight of **150** gr and a **standard deviation** of **15** gr. Make a 95% confidence interval for μ.

Solution

The IC formula for X normally distributed and σ NOT known is:

$$(1-\alpha)\%IC(\mu) = \left[\bar{x} - t_{n-1;\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}\,;\,\bar{x} + t_{n-1;\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}\right]$$

We have: $\bar{x}$ = 150 , n=30 , $\hat{\sigma}$ = 15

The standard error is: $\sigma_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}}$ = 15/√30 = 15 / 5.48 = 2.74

The degrees of freedom for the calculation of the Student's T are n-1=29 and $(1-\alpha) =$ 0.95, so we have that: $t_{n-1;\alpha/2}$= $t_{(29;0.025)}$ that from the table of the Student's T is: $t_{(29;0.025)}$ =2.045

95% $IC(\mu)$ = [$\bar{x}$ - $t_{n-1;\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$; $\bar{x}$ + $t_{n-1;\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$] = [150 – 2.045 · 2.74 ; 150 + 2.045 · 2.74] =

= [144.4 ; 155.6]

Exercise 4

For a data set obtained from a sample, n = **25** and $\bar{x}$ = **35**. The sample standard deviation is **3.1**. Make a **95%** confidence interval for μ.

Solution

$\bar{x}$ = 35 , n=25 , $\hat{\sigma}$ = 3.1

$\sigma_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}}$ = 3.1/√25 = 3.1 / 5 = 0.62

Degrees of freedom=n-1=24, so we have that, $t_{(24;0.025)}$=2.064

95% confidence interval: [$\bar{x}$ - 2.064 $\hat{\sigma}$/√n ; $\bar{x}$ + 2.064 $\hat{\sigma}$/√n] = [35 – 2.064 · 0.62 ; 35 + 2.064 · 0.62] =

= [33.72 ; 36.28]

Exercise 5

A sample of **30** products has a **mean** weight of **150** gr and a **standard deviation** of **15** gr. Make a **95%** confidence interval for μ.

Solution

$\bar{x}$ = 150 , n=30 , s = 15

$\sigma_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}}$ = 15/√30 = 15 / 5.48 = 2.74

Degrees of freedom=n-1=29, so we have that, $t_{(29;0.025)}$=2.045

95% confidence interval: [$\bar{x}$ - 2.045 $\hat{\sigma}$/√n ; $\bar{x}$ + 2.045 $\hat{\sigma}$/√n] = [150 – 2.045 · 2.74 ; 150 + 2.045 · 2.74] =

= [144.4 ; 155.6]

Confidence Interval for proportion p

The (1-α)% ***confidence interval for p*** is:

Mean of Bernoullian distribution

$$\hat{p} \pm z\sqrt{\frac{\hat{p}\cdot(1-\hat{p})}{n}}$$

Variance of Bernoullian distribution

Exercise 6

By a sample of **100** customers, **10** reply that are satisfied with the service they received. Calculate a **90**% confidence interval for the proportion of satisfied customers.

Solution

The IC formula for X Bernoullian r.v. and π unknown is:

$$(1-\alpha)\%IC(\pi) \cong \left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}\,;\hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}\right]$$

The $z_{\alpha/2}$ for $(1-\alpha) = 0.90$ is $z_{\alpha/2} = z_{0.05} = 1.645$

The observed value (proportion of the sample) is:

$$\hat{p} = \frac{n.\,customers\ satified}{n.\,customers} = \frac{10}{100} = 0.10 = 10\%$$

The variance of a Bernoullian variable $\hat{p}$ is $\hat{p}\ \cdot (1-\hat{p})$

The standard error of $\hat{p}$ is equal to $\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} = \sqrt{\frac{0.10\ 0.90}{100}} = 0.03$

90% $IC(\pi) \cong [\hat{p}$ - $z_{\alpha/2}$ √ p(1- p)/n ; $\hat{p}$ + $z_{\alpha/2}$√ p(1- p)/n] = [0.1 – 1.64 · 0.03 ; 0.1 + 1.64 · 0.03] =

= [0.0508, 0.1492]

Exercise 7

Sample of **80** customers **60** reply they are satisfied with the service they received. Calculate a **95%** confidence interval for the proportion of satisfied customers.

Solution

The observed value (proportion of the sample) is:

$$\hat{p} = \frac{n.\,customers\ satified}{n.\,customers} = \frac{60}{80} = 0.75 = 75\%$$

Variance of $\hat{p}$: $\frac{Var(X)}{n} = \frac{\hat{p}\,(1-\hat{p})}{n} = 0.0023$

Standard deviation of $\hat{p}$: $\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}$ = 0.048

95% Confidence interval for the true proportion π :

95% $IC(\pi) \cong [\hat{p}$ - $z_{\alpha/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}$; $\hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}] = [0.75 - 1.96 \cdot 0.048\,;\ 0.75 + 1.96 \cdot 0.048] =$

= [0.654, 0.846]

THEME #2

—

Confidence intervals on two means

Confidence intervals on two means with independent samples

σ **known** and the **same** for the two means

$$\bar{x}_1 - \bar{x}_2 \pm z\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

σ **known** and **different** for the two means σ_1 and σ_2

$$\bar{x}_1 - \bar{x}_2 \pm z\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

σ **unknown**
(σ_1 and σ_2 **unequal**)

$$\bar{x}_1 - \bar{x}_2 \pm t\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

with df = $\dfrac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}$

σ **unknown**
(σ_1 and σ_2 **equal**) ***pooled t procedure*** (*df*=n_1+n_2-2)

$$\bar{x}_1 - \bar{x}_2 \pm t s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

Exercise 1

A sample of **40** pine trees grown on the north side of a hill has a **mean** of **25.4** metres and a **standard deviation** of **2.1** metres.
A second sample of **40** trees from the south side has a **mean** of **23.2** metres and **a standard deviation** of **1.7** metres.
Find the **95%** confidence interval for the difference in the mean heights of the two populations of trees.

Solution

$$(1-\alpha)\%IC(\mu_1-\mu_2)=\bar{x}_1-\bar{x}_2\pm t\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

$\bar{x}_1 = 25.4$ with $s_1 = 2.1$ and $\bar{x}_2 = 23.2$ with $s_2 = 1.7$

$s_1^2 = 4.41$ and $s_2^2 = 2.89$

$$SE = \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} = \sqrt{\frac{4.41}{40}+\frac{2.89}{40}} = \sqrt{0.1825} = 0.427$$

Exercise 2

A sample of 40 pine trees grown on the north side of a hill has a mean of 25.4 metres and a standard deviation of 2.1 metres.
A second sample of 40 trees from the south side has a mean of 23.2 metres and a standard deviation of 1.7 metres.
Find the 95% confidence interval for the difference in the mean heights of the two populations of trees.

Solution

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} = \frac{0.1825^2}{\frac{(\frac{4.41}{40})^2}{39} + \frac{(\frac{2.89}{40})^2}{39}} = \frac{0.033}{0.0003 + 0.0001} = 82.5 \cong 83$$

$$t_{83;0.025} = 2.884$$

So,

$$95\% IC(\mu_1 - \mu_2) = \bar{x}_1 - \bar{x}_2 \pm t\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (25.4 - 23.2) \pm 2.884 \cdot 0.427 = (0.97; 3.43)$$