Estimation of intervals

- How do we estimate the average height of an 18 year old males ?
- Take a sample.
- Estimate sample mean $\overline{x}$.
- How sure can we about our estimate ?

Point and interval estimation

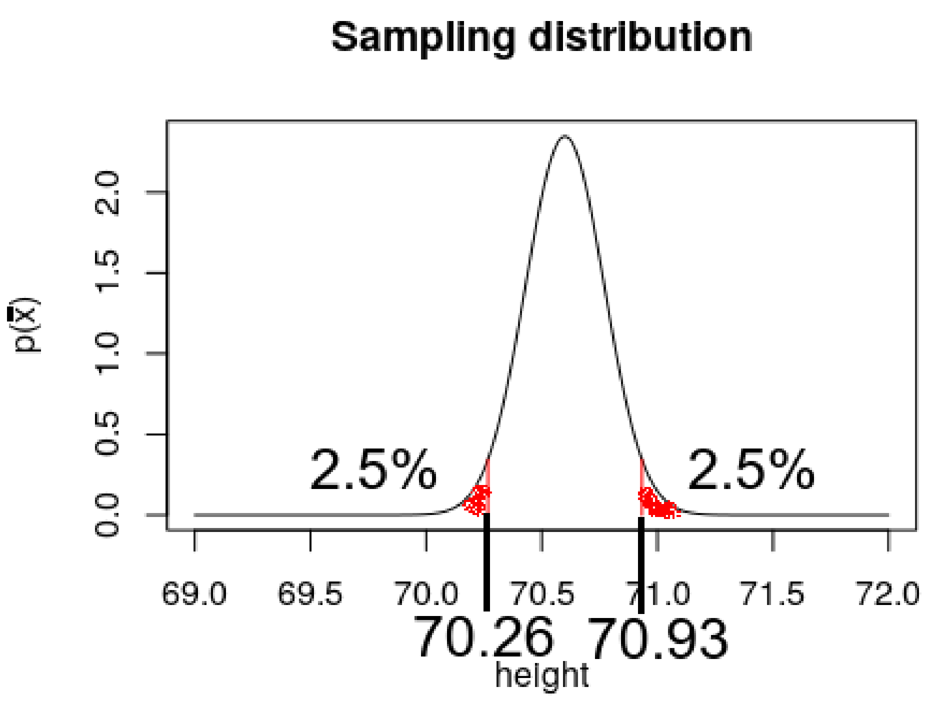
- Assuming this single number $\bar{x}$ as population mean, is called a **point estimate**. It gives no indication of how reliable the estimate is.
- In contrast, an **interval estimation** states the **margin of error** E. The estimate takes the form $[\bar{x} - E, \bar{x} + E]$, which states that a certain proportion, say 95%, of the means estimated from sample data fall in this interval. Such an interval is called a 95% confidence interval for μ.

Example height estimation

- We have sample n=100 of height measure from men aged 18. The sample mean is 70.6 inches and the standard deviation is s=1.7.
- We know how the sample is liked to the sample statistics. Therefore, the mean of the sample distribution is

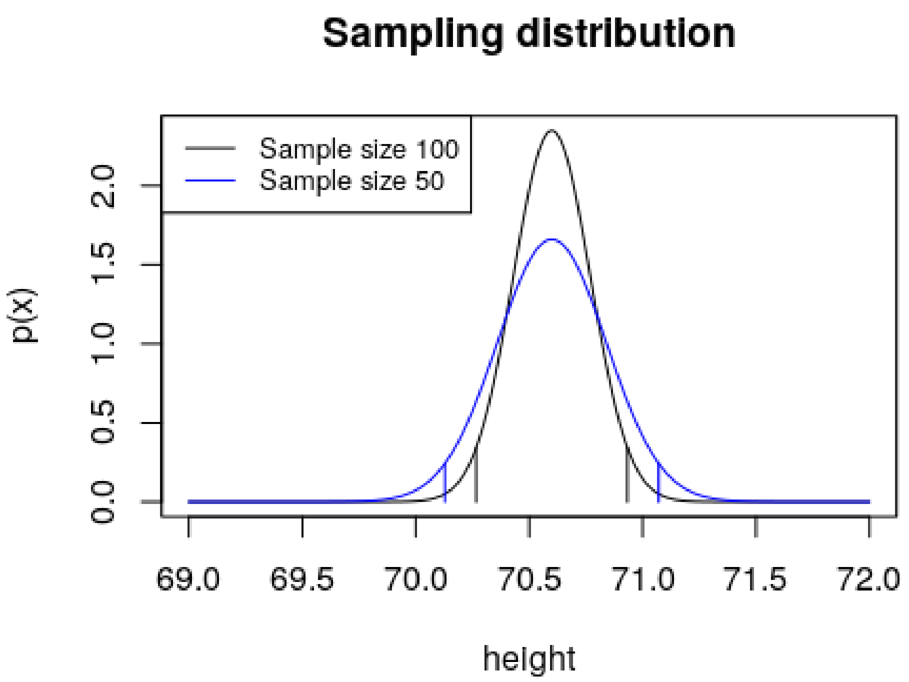
$\mu_{\bar{x}} = \mu = 70.6 \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.7}{10} = 0.17$

- $\bar{X}$ = [70.26,70.93]



How about the sample size

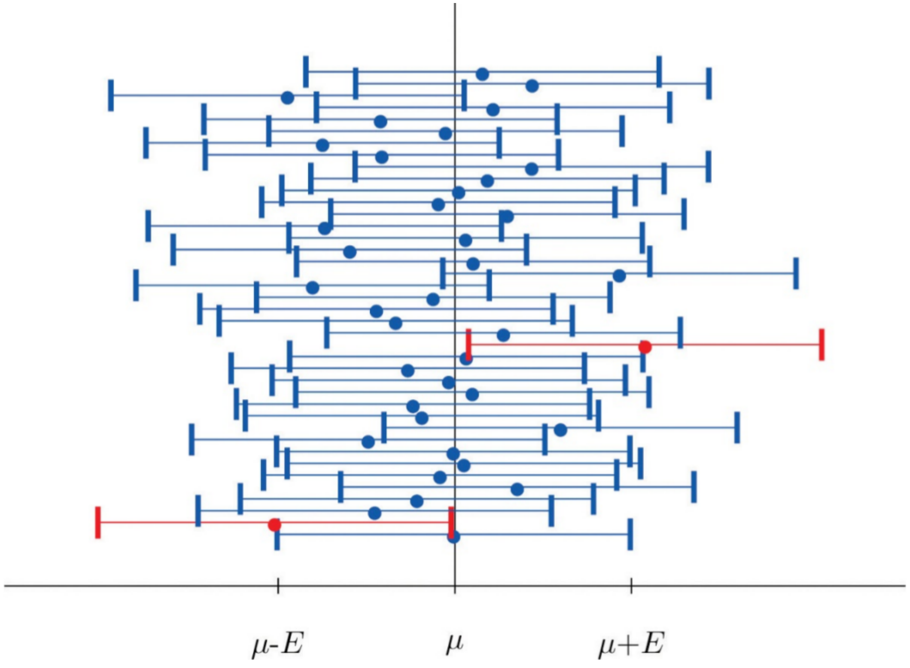
- If we had the same sample mean and standard deviation, but a sample size of 50, the estimate of the standard deviation of the sample mean would be higher. And our estimate changes accordingly.
- $\mu_{\bar{x}} = \mu = 70.6$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.7}{\sqrt{(50)}} = 0.24$
- $\bar{X} = [70.13, 71.07]$



Confidence interval

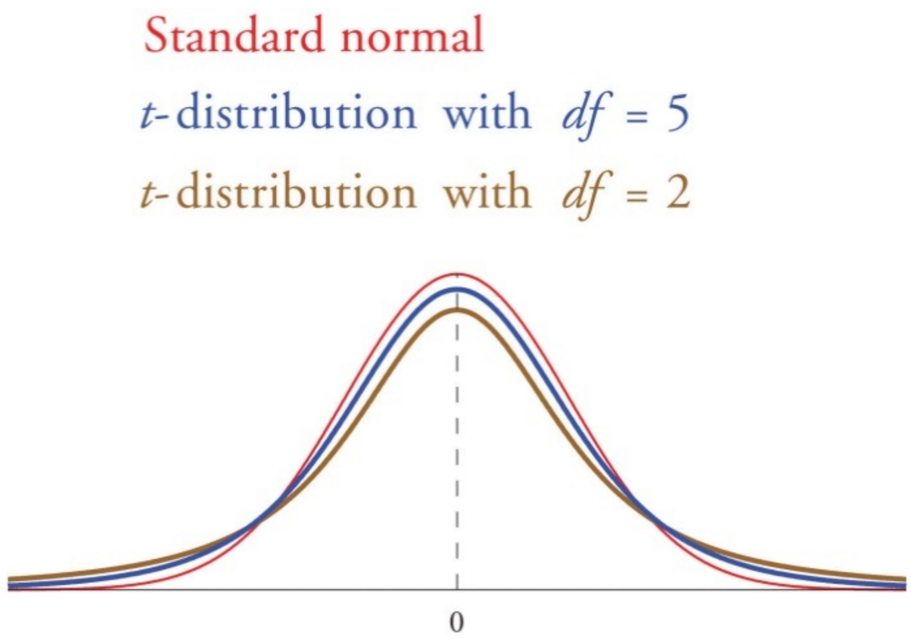
- We construct a 95% confidence interval E, than 95% of the estimated sample means will lie within the interval [μ+E, μ-E].

- With a 95% confidence interval estimated from a sample, we would expect 5% of drawn samples to have a $\overline{X}$ outside the confidence interval.
- With 40 drawn samples, we would expect 40*0.05= 2 samples to have a $\overline{X}$ outside the interval.



Small sample sizes

- The Central Limit Theorem applies for samples sizes n>=30. For smaller samples sizes, this does not apply. However, if the population distribution is normal and we know σ, we can still use the previous estimation.
- If the population is normal distributed, the standard deviation σ unknown and the sample size is small, we use the Student's t-distribution with n-1 degrees of freedom.



Example

- A sample of n=15 drawn from a normal distributed population has a sample mean 35 and standard deviation 14. Construct a 95% confidence interval for the population mean.
- $\bar{X}$ =[27.24,42.75]