Confidence Interval for a Proportion

Confidence Interval = p +/- z*(√p(1-p) / n)

where: p: sample proportion z: the chosen z-value n: sample size

A random sample of 500 apples was taken from, a large consignment and 60 were found to be bad. Obtain the 95%, 98% confidence limits for the percentage number of bad apple in the consignment.

```
library(glue)
n<-500;
p<-(60/500)
SE <- sqrt(p*(1-p)/n)
z_star<- qnorm(1-(1 - 0.95)/2)
ME<-z_star*SE
glue("({p - ME}, {p + ME})")
```

```
library(glue)
n<-500;
p<-(60/500)
SE <- sqrt(p*(1-p)/n)
z_star<- qnorm(1-(1 - 0.98)/2)
ME<-z_star*SE
glue("({p - ME}, {p + ME})")
```

A sample of 900 members has a mean 3·4 cms, and s.d. 2·61 cms. If the population is normal and its mean is unknown, find the 95% and 98'% fiucial limits of true mean.

Xbar +/- z * sigma/sqrt(n)

```
library(glue)
n<-900;
sigma<-2.61
Xbar<-3.4
SE <- sigma/sqrt(n)
z_star<- qnorm(1-(1 - 0.95)/2)
ME<-z_star*SE
glue("({Xbar - ME}, { Xbar + ME})")
```

```
n<-900;
sigma<-2.61
Xbar<-3.4
SE <- sigma/sqrt(n)
z_star<- qnorm(1-(1 - 0.98)/2)
ME<-z_star*SE
glue("({Xbar - ME}, { Xbar + ME})")
```

solve

The mean muscular endurance score of a random sample of 60 subjects was found to be 145 with a s.d. of 40. Construct a 95% confidence interval for the true mean. Assume the sample size to be large enough for normal approximation.

Confidence Interval for a Difference in Proportions

Confidence interval = (p1–p2) +/- z*√(p1(1-p1)/n1 + p2(1-p2)/n2)

where:

p1, p2: sample 1 proportion, sample 2 proportion

z: the z-critical value based on the confidence level

n1, n2: sample 1 size, sample 2 size

Write R code for the above formula and solve the following problem

A medical researcher conjectures that smoking can result in the wrinkled skin around the eyes. The researcher recruited **150 smokers** and **250 nonsmokers** to take part in an observational study and found that **95** of the **smokers** and **105** of the **nonsmokers** were seen to have prominent wrinkles around the eyes (based on a standardized wrinkle score administered by a person who did not know if the subject smoked or not). Find **CI** for the true difference that would exist between these two groups in the population.

Confidence Interval for a Difference in Means

Then the confidence interval for the difference in the two population means is

$$\bar{x}_1 - \bar{x}_2 \pm (z \text{ critical value})\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

$$S.E.\ (\bar{x}_1 - \bar{x}_2) = \sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)} = \sqrt{(s_1^2/n_1) + (s_2^2/n_2)}$$

Write R code for the above formula and solve the following problem

In a certain factory there are two independent processes manufacturing the same item. The average weight in a sample of 250 items produced from one process is found to be 120 ozs. with a standard deviation of 12 ozs. while the corresponding figures in a sample of 400 items from ihe other process are 124 and 14. Obtain the standard error of difference between the two. sample means. Find the 99% confidence limits for the difference in the average weights of items produced by the two processes respectively.

Answer

$$
\begin{aligned}
|\bar{x}_1 - \bar{x}_2| \pm 2{\cdot}58\ S.E.\ (\bar{x}_1 - \bar{x}_2) &= 4 \pm 2{\cdot}58 \times 1{\cdot}034 \\
&= 4 \pm 2{\cdot}67 \text{ (approx.)} = 6{\cdot}67 \text{ and } 1{\cdot}33 \\
1{\cdot}33 < |\mu_1 - \mu_2| &< 6{\cdot}67
\end{aligned}
$$

P-value from z score

Right-tailed test

Suppose we want to find the p-value associated with a z-score of 2.02 in a right-tailed hypothesis test.

P = P(Z > 2.02) = 0.0217

pnorm(q=2.02, lower.tail=FALSE)

[1] 0.02169169

Left-tailed test

Suppose we want to find the p-value associated with a z-score of -0.77 in a left-tailed hypothesis test.

pnorm(q=-0.77, lower.tail=TRUE)

[1] 0.2206499

Two-tailed test

Suppose we want to find the p-value associated with a z-score of 2.83 in a two-tailed hypothesis test.

P = P(|Z| > 2.83) = 2P(Z < −2.83) = 0.0046 2*pnorm(q=2.83, lower.tail=FALSE)

[1] 0.0046548

T test (two tiled)

P = P(|T| > 2.06) ≈ 0.06. 2*pt(q=2.06, 14,lower.tail=FALSE)

[1] 0.05849421

* here 14 is degrees of freedom

Similarly you can do for one tailed test (t)