Significance (or hypothesis) tests
Quantiles

Def. Let f be a density function and $n \in \mathbb{N}$. The values $\xi_1, \dots, \xi_{n-1}$ determined such that

$$\int_{-\infty}^{\xi_1} f(x)\,dx = \int_{\xi_1}^{\xi_2} f(x)\,dx = \dots = \int_{\xi_{n-1}}^{\infty} f(x)\,dx = \frac{1}{n} \overset{\text{not}}{=} \alpha$$

are called quantiles or α-quantiles.

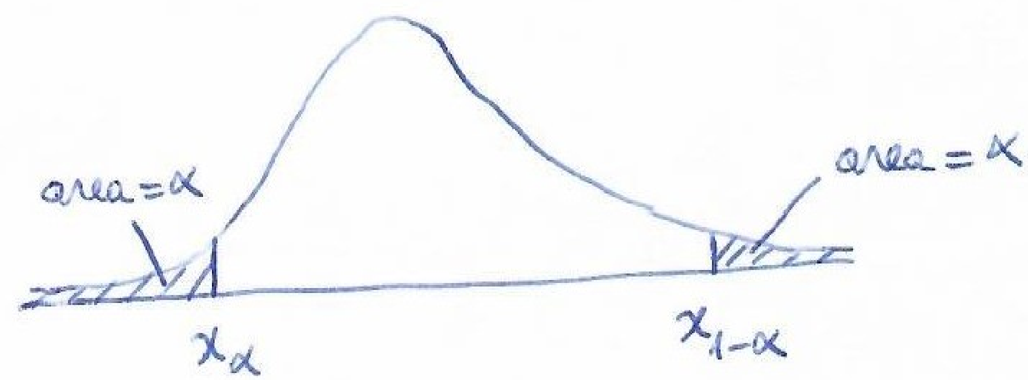
In the special cases where $\alpha = \frac{1}{2}, \frac{1}{4}, \frac{1}{10}, \frac{1}{100}$, these α-quantiles are called median, quartiles, deciles, percentiles.

Quantiles

ξ_1 is the inferior α-quantile $\overset{\text{not}}{=} x_\alpha$

ξ_{n-1} is the superior α-quantile $= x_{1-\alpha}$

$F(\xi_1) = 1 - F(\xi_{n-1}) = \alpha$, where F is the distribution function



X^2, t and F distributions

1) χ^2 (chi-square) with k degrees of freedom

$$\chi^2_k = \sum_{i=1}^{k} U_i^2 \quad , \text{ where } U_i \sim \text{iid } N(0,1) \quad i=\overline{1,k}$$

The distribution function is

$$H_k(x) = \frac{1}{2^{\frac{k}{2}} \cdot \Gamma\left(\frac{k}{2}\right)} \int_0^x u^{\frac{k}{2}-1} \cdot e^{-\frac{u}{2}} \, du ,$$

where $\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} \cdot e^{-u} \, du$

X^2, t and F distributions

2) t (or Student's) distribution with k degrees of freedom

$$t_k = \frac{U}{\sqrt{\frac{\chi_k^2}{k}}} \, , \quad U \sim N(0,1) \text{ independent of } \chi_k^2$$

The distribution function is

$$G_k(z) = \frac{1}{\sqrt{\pi k}} \, \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} \int_{-\infty}^{z} \left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}} dx$$

Its α-quantiles satisfy the relation

$$t_{k,\alpha} = - t_{k,1-\alpha}$$

χ^2, t and F distributions

3) F (or Snedecor's) distribution

$$F_{k_1,k_2} = \frac{k_2 \cdot \chi^2_{k_1}}{k_1 \cdot \chi^2_{k_2}} \quad , \quad \chi^2_{k_1} \text{ and } \chi^2_{k_2} \text{ are independent}$$

The distribution function is

$$S_{k_1,k_2}(x) = \left(\frac{k_1}{k_2}\right)^{\frac{k_1}{2}} \frac{\Gamma\left(\frac{k_1+k_2}{2}\right)}{\Gamma\left(\frac{k_1}{2}\right)\Gamma\left(\frac{k_2}{2}\right)} \int_0^x z^{\frac{k_1}{2}-1} \cdot \left(1+\frac{k_1}{k_2}\cdot z\right)^{-\frac{k_1+k_2}{2}} dz$$

Significance tests

They are statistical tools to measure whether a set of observed data produces a result more extreme than what chance might produce. if the result is outside the limits of chance variation, it is said to be statistically significant.

Suppose that $\{X_1, \dots, X_n\}$ are iid as the random variable X. We are interested to check whether X satisfies the hypothesis H_0 (called null hypothesis). against H_A (called alternative hypothesis). A test is a criterion (a rule) based upon $X_1, \dots, X_n$.

Significance tests

A classical procedure to build a test is as follows:

- the sample space S^n is divided in $W \cup \overline{W} = S^n$

- if the statistical data $(x_1, \ldots, x_n)' \in \overline{W}$ then H_0 is rejected
  if $(x_1, \ldots, x_n)' \in W$ then H_0 is accepted

- $\overline{W}$ is called the critical region of the test and it is determined such that $P\left((X_1, \ldots, X_n)' \in \overline{W} \mid H_0 \text{ is true}\right) = \alpha$, where α is a given value (usually 0.05 or 0.01) called significance level.

- W is called the acceptance region of the test

Significance tests

Decision \ True hypothesis	H_0	H_A
H_0	$(H_0 \mid H_0)$	$(H_0 \mid H_A)$
H_A	$(H_A \mid H_0)$	$(H_A \mid H_A)$

Type II error = failure to reject a false null hypothesis

Type I error = rejection of a true null hypothesis

$\alpha = P(H_A \mid H_0)$ - risk of type I

$\beta = P(H_0 \mid H_A)$ - risk of type II

$\Pi = 1 - \beta = P(H_A \mid H_A)$ - power of the test

We are interested in tests with α, β small, Π big.

Confidence intervals

Let X be a random variable with the distribution F_θ, $\theta \in \Theta \subseteq \mathbb{R}$ and $X_1, \dots, X_n$ a random sample i.i.d as X.
if $\alpha \in (0,1)$ (small) and $A_\alpha, B_\alpha : S^n \to \mathbb{R}$ are two measurable function such that:

i) $A_\alpha(x_1, \dots, x_n) \leq B_\alpha(x_1, \dots, x_n), \quad \forall (x_1, \dots x_n)' \in S^n$

ii) $P_\theta\left(A_\alpha(X_1, \dots, X_n) \leq \theta \leq B_\alpha(X_1, \dots, X_n)\right) = 1-\alpha$

then, for the statistical data $(x_1, \dots, x_n)$, the interval

$C_{n;1-\alpha}(x_1, \dots, x_n) = \left[A_\alpha(x_1, \dots x_n), B_\alpha(x_1, \dots, x_n)\right]$ is called confidence interval for θ, with the confidence level $1-\alpha$.

Confidence intervals

How to determine confidence intervals – the method of the pivotal function

Def. $g: S^n \times \Theta \to \mathbb{R}$ is called pivotal function iff:

- $g((x_1, \dots, x_n), \theta)$ is continuous and strictly monotone with regard to θ
- $g((X_1, \dots, X_n), \theta)$ has the distribution independent of θ

Given $\alpha \in (0,1)$, we determine two quantiles $a(\alpha), b(\alpha)$ of the distribution of g, such that

$$P_\theta\left(a(\alpha) \le g((X_1, \dots, X_n), \theta) \le b(\alpha)\right) = 1-\alpha$$

Confidence intervals

The confidence interval $C_{n;1-\alpha}(x_1,\dots,x_n) = [A_\alpha(x_1,\dots,x_n), B_\alpha(x_1,\dots,x_n)]$
is determined by solving the two inequations

$$a(\alpha) \le g((X_1,\dots,X_n),\theta)$$

$$g((X_1,\dots,X_n),\theta) \le b(\alpha)$$

Obs. The quantiles $a(\alpha)$ and $b(\alpha)$ are not uniquely determined, therefore the confidence interval is not unique. We are interested to determine the shortest confidence interval, given a confidence level.

Confidence intervals

Example $\quad X \sim N(\mu, \sigma^2) \quad$ μ unknown; σ^2 known

$X_1, \dots, X_n$ a random sample for X

$$\left[\begin{array}{l} Y_i \sim N(\mu_i, \sigma_i^2) \text{ indep.} \\ Y = a_1 Y_1 + \dots + a_n Y_n \end{array} \Rightarrow Y \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)\right]$$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \Rightarrow \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

We take the pivotal function

$$g(X_1, \dots, X_n; \mu) = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

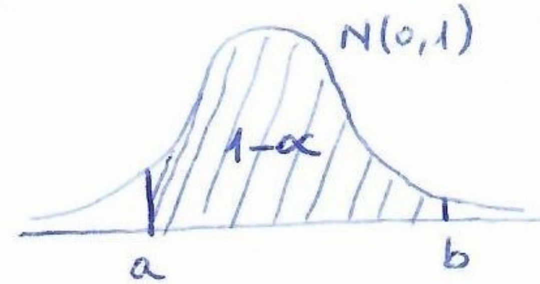
does not depend on μ

Confidence intervals

g is strictly decreasing with regard to μ

For a given $\alpha \in (0,1)$, let a, b be two quantiles of $N(0,1)$ such that

$$P\left(a \leq \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq b\right) = 1-\alpha$$



(a,b) can be asymmetric

$x_1, \dots, x_n$ - statistical data

$$a \leq \frac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq b \iff \overline{x} - b\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{x} - a\frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow C_{n;1-\alpha} = \left[\overline{x} - b\frac{\sigma}{\sqrt{n}},\ \overline{x} - a\cdot\frac{\sigma}{\sqrt{n}}\right]$$

Confidence intervals

We are interested in finding the shortest interval

$l = \frac{\sigma}{\sqrt{n}}(b-a)$ the length of $C_{n;1-\alpha}$

b depends on a — $b = b(a)$

$$\begin{cases} F_{N(0,1)}(b) - F_{N(0,1)}(a) = 1-\alpha, \\ \min\left\{\frac{\sigma}{\sqrt{n}}(b-a)\right\} \end{cases}$$

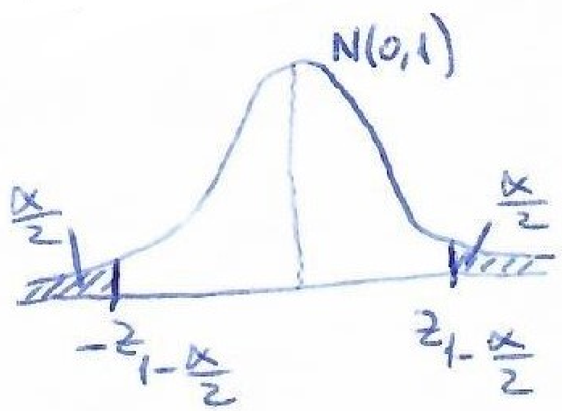
$F_{N(0,1)}$ = distribution function of $N(0,1)$

$$\begin{cases} f_{N(0,1)}(b) \cdot \frac{\partial b}{\partial a} - f_{N(0,1)}(a) = 0 \\ \frac{\partial b}{\partial a} - 1 = 0 \end{cases}$$

Confidence intervals

$\Rightarrow f_{N(0,1)}(b) = f_{N(0,1)}(a) \Rightarrow$ the shortest confidence interval is the symmetric one

$b = z_{1-\frac{\alpha}{2}} \qquad a = -z_{1-\frac{\alpha}{2}}$



$$\Rightarrow C^{*}_{n;1-\alpha}(x_1,\dots,x_n) = \left[\bar{x} - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

is the shortest confidence interval for μ, with the confidence level $1-\alpha$

if $\alpha = 0.05 \Rightarrow z_{1-\frac{\alpha}{2}} = 1.96$

Significance tests

How to build a significance test – the intuitive method based upon acceptance regions

$H_0: \{\theta = \theta_0\}$ simple null hypothesis

$H_A: \{\theta \neq \theta_0\}$ composite alternative hypothesis (can have many different θs)

α is the significance level (given)

Suppose that there is a pivotal function g, then we have two quantiles $a(\alpha)$ and $b(\alpha)$ such that:

$$P_{\theta_0}\left(a(\alpha) \leq g((X_1, \dots, X_n); \theta_0) \leq b(\alpha)\right) = 1-\alpha$$

Significance tests

We set the acceptance region of H_0 at the significance level α as follows:

$$W_{n;1-\alpha}(\theta_0) = \{(x_1, \dots, x_n)' \mid a(\alpha) \leq g((x_1, \dots, x_n); \theta_0) \leq b(\alpha)\}$$

and the critical region as complement $\overline{W}_{n;1-\alpha}(\theta_0)$.

The probability of type I error is

$$P(H_A \mid H_0) = P_{\theta_0}\left((X_1, \dots, X_n) \in \overline{W}_{n;1-\alpha}(\theta_0)\right) = 1 - (1-\alpha) = \alpha$$

Significance tests

Most applications of hypothesis testing control the probability of making a type I error, but it is not always the case with the type II error. With this regard, an operating characteristic curve can be plotted to show how the sample size affects the probability of making a type II error.

The operating characteristic function of H_0 is defined as follows:

$$OC(\theta) = P_\theta\left((X_1, \dots, X_n) \in W_{n;1-\alpha}(\theta_0)\right) \qquad \left(= P(H_0 | H_A)\right)$$

Usually, the smaller the difference between θ and θ_0, the larger the sample size n needs to be in order to get a small β.

Significance tests

For example, if the true θ is 1.5 and we test $H_0: \{\theta = 1.7\}$, the probability of accepting H_0 when H_A is true is very high. if we increase the sample size n, the probability for the test to perceive the difference (i.e. H_0 is rejected) increases.

On the other hand, if we test $H_0: \{\theta = 2.5\}$, we expect a much lower probability of accepting H_0.

One sample t-test

- for the mean of a normal distribution with unknown variance.

$X \sim N(\mu, \sigma^2)$, σ^2 unknown

$X_1, \dots, X_n$ a random sample $N(\mu, \sigma^2)$

$$H_0: \{\mu = \mu_0\} \qquad H_A: \{\mu \neq \mu_0\}$$

One sample t-test

I First, we determine the confidence interval for μ, with the confidence level $1-\alpha$.

Recall that $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is MLE for μ - it is unbiased

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ is unbiased for σ^2

it can be proven that $\frac{n-1}{\sigma^2}S^2 = \sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$ and

$\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}$ is independent from $\frac{n-1}{\sigma^2}S^2$ (for proof, see Dumitrescu & Bătătorescu, pag 176)

it follows that $\dfrac{\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}}{\sqrt{\frac{n-1}{\sigma^2}\cdot S^2\cdot\frac{1}{n-1}}} = \frac{\sqrt{n}(\overline{X}-\mu)}{S} \sim t_{n-1}$

One sample t-test

Now, we take the pivotal function $g((x_1, \dots, x_n); \mu) = \frac{\sqrt{n}(\bar{x} - \mu)}{s}$

g is strictly decreasing in μ and its distribution t_{n-1} does not depend on μ.

For $\alpha \in (0,1)$, there are two quantiles $a(\alpha), b(\alpha)$ for the distribution t_{n-1} such that

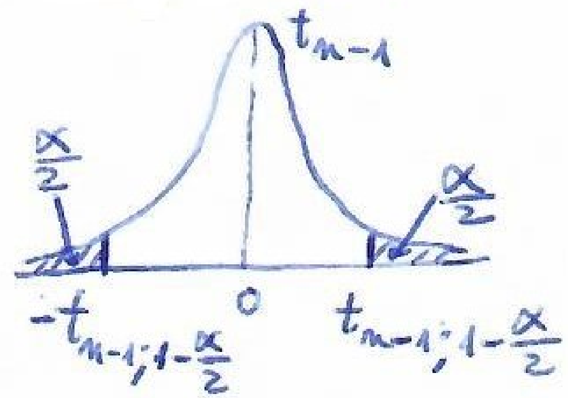
$$P_\mu\left(a(\alpha) \leq \frac{\sqrt{n}(\bar{x} - \mu)}{s} \leq b(\alpha)\right) = 1 - \alpha$$

One sample t-test

Thus, $C_{n;1-\alpha}(x_1, \ldots, x_n) = \left[\bar{x} - b\frac{s}{\sqrt{n}}, \bar{x} - a \cdot \frac{s}{\sqrt{n}}\right]$ is a confidence interval for μ, with the confidence level $1-\alpha$.

The shortest confidence interval is the symmetric one, where

$$b = t_{n-1;1-\frac{\alpha}{2}} \qquad a = -t_{n-1;1-\frac{\alpha}{2}}$$



One sample t-test

II Now, we build the one-sample T-test

The acceptance region for $H_0: \{\mu = \mu_0\}$ is the following:

$$W_{n;1-\alpha}(\mu_0) = \left\{ (x_1, \dots, x_n) \,\middle|\, -t_{n-1;1-\frac{\alpha}{2}} \le \frac{\sqrt{n}(\bar{x} - \mu_0)}{s} \le t_{n-1;1-\frac{\alpha}{2}} \right\}$$

$$= \left\{ (x_1, \dots, x_n) \,\middle|\, \mu_0 - t_{n-1;1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \le \bar{x} \le \mu_0 + t_{n-1;1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \right\}$$

The critical region is the complement $\overline{W_{n;1-\alpha}(\mu_0)}$.

One sample t-test

The operating characteristic function

$$OC(\mu) = P_\mu\left(-t_{n-1;1-\frac{\alpha}{2}} \le \frac{\sqrt{n}(\overline{X}-\mu_0)}{S} \le t_{n-1;1-\frac{\alpha}{2}}\right)$$

$$= P_\mu\left(-t_{n-1;1-\frac{\alpha}{2}} \le \frac{\sqrt{n}(\overline{X}-\mu)}{S} + \frac{\sqrt{n}(\mu-\mu_0)}{S} \le t_{n-1;1-\frac{\alpha}{2}}\right)$$

$$= P_\mu\left(-t_{n-1;1-\frac{\alpha}{2}} - \frac{\sqrt{n}(\mu-\mu_0)}{S} \le \frac{\sqrt{n}(\overline{X}-\mu)}{S} \le t_{n-1;1-\frac{\alpha}{2}} - \frac{\sqrt{n}(\mu-\mu_0)}{S}\right)$$

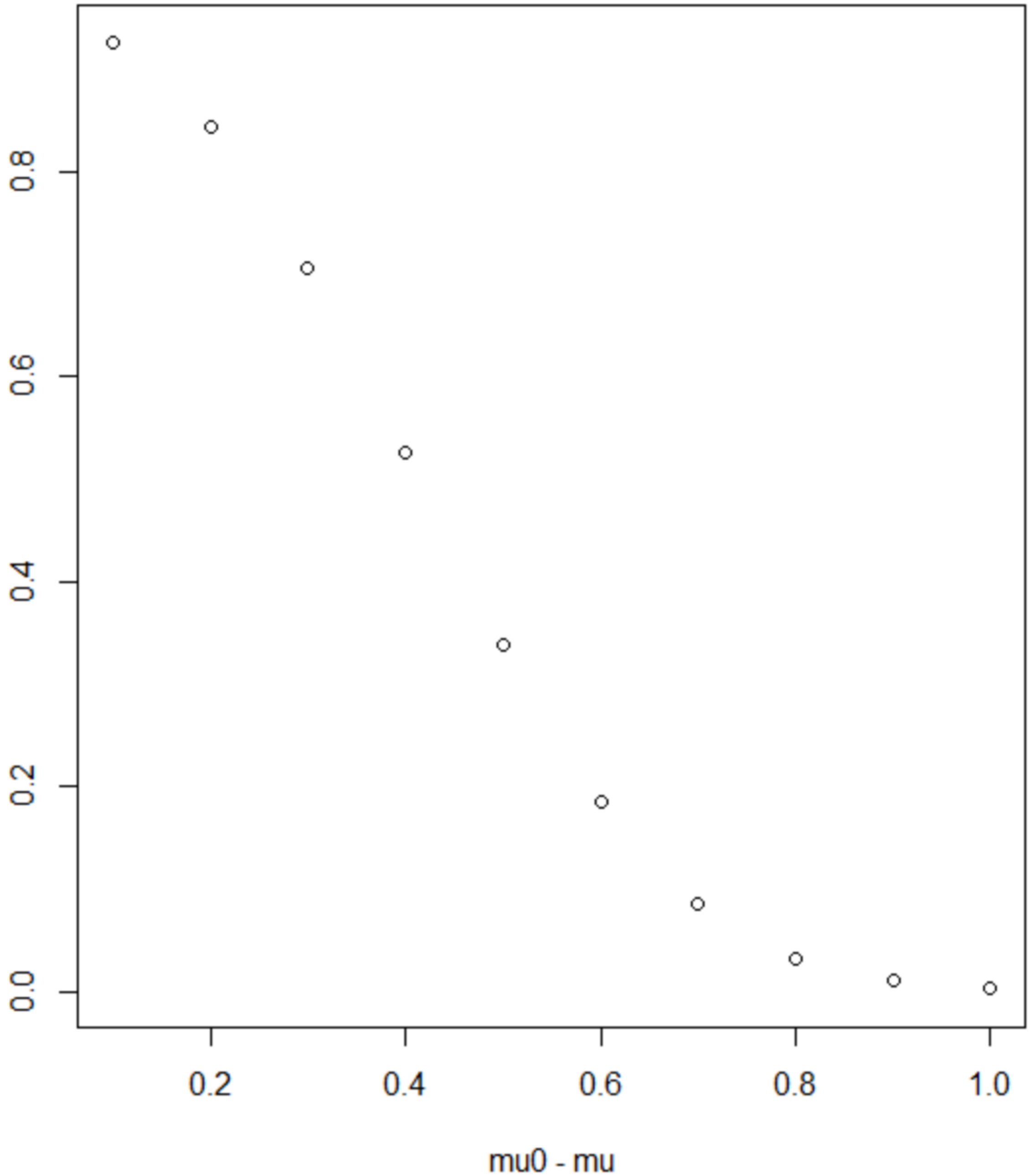
$$= G_{n-1}\left(t_{n-1;1-\frac{\alpha}{2}} - \frac{\sqrt{n}(\mu-\mu_0)}{S}\right) - G_{n-1}\left(-t_{n-1;1-\frac{\alpha}{2}} - \frac{\sqrt{n}(\mu-\mu_0)}{S}\right)$$

G_{n-1} is the distribution function for t_{n-1}

[if a random variable $X \sim F(x)$
then $P(a \le X \le b) = F(b) - F(a)$]

One sample t-test – the OC function

```
mu=2  #the true value of the parameter
mu0=seq(2.1,by=0.1,3)
sample size = 100
```




```
mu=2  #the true value of the parameter
mu0=2.1
sample size = 200,300, ...,1000
```