

Statistical inference I and II

Statistical inference

Stanislav Katina

ÚMS, MU

13.03.2024 06:03

Book Aplikovaná štatistická inferencia I

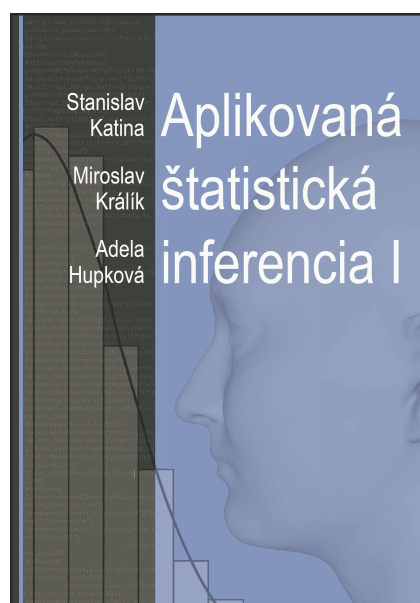


Figure 1: Book Aplikovaná štatistická inferencia I

- 1 Overview of testing of statistical hypotheses
- 2 Three test statistics
- 3 Confidence intervals
- 4 Overview of the tests
- 5 Generalised hypotheses
- 6 One-sample tests about μ
- 7 Paired tests about μ
- 8 One-sample tests about σ^2
- 9 One-sample tests about skewness and kurtosis
- 10 One-sample tests about correlation coefficient

6.5 One-sample tests about μ (cont.)

H_{01} vs H_{11} . Let $X \sim N(\mu, \sigma^2)$, where $\boldsymbol{\theta} = (\mu, \sigma^2)^T$. Then the **likelihood function** is equal to

$$L(\boldsymbol{\theta}|\mathbf{x}) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right)$$

and the **log-likelihood function**

$$\ell(\boldsymbol{\theta}|\mathbf{x}) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2.$$

If H_{01} is true, $\hat{\boldsymbol{\theta}}_0 = (\theta_0, \hat{\theta}_{2|0})^T$, where $\theta_0 = \mu_0$ and

$\hat{\theta}_{2|0} = \hat{\sigma}_0^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_0)^2 = s_n^2 + (\bar{x} - \mu_0)^2$, where $s_n^2 = \hat{\sigma}^2$. The variance $\hat{\sigma}_0^2$ is the solution of

$$S_2(\boldsymbol{\theta}_0) = \frac{\partial}{\partial \sigma^2} \ell(\boldsymbol{\theta}|\mathbf{x}),$$

where using $\mu = \mu_0$, $\sigma^2 = \hat{\sigma}_0^2$ and

$$\frac{1}{2} \left(\frac{1}{\hat{\sigma}_0^4} \sum_{i=1}^n (x_i - \mu_0)^2 - \frac{n}{\hat{\sigma}_0^2} \right) = 0.$$

we get $\hat{\sigma}_0^2$.

6.6 One-sample tests about μ (cont.)

If H_{01} is not true, MLE of θ is then equal to

$$\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)^T = (\bar{x}, \hat{\sigma}^2)^T = \left(\frac{1}{n} \sum_{i=1}^n x_i, \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^T.$$

Then

$$\begin{aligned} u_{\text{LR}} &= -2 \ln(\lambda(\mathbf{x})) \\ &= 2(\ell(\hat{\theta}|\mathbf{x}) - \ell(\theta_0|\mathbf{x})) \\ &= 2 \left(\left(-\frac{n}{2} (\ln(2\pi) + \ln \hat{\sigma}^2 + 1) \right) - \left(-\frac{n}{2} (\ln(2\pi) + \ln \hat{\sigma}_0^2 + 1) \right) \right) \\ &= n \ln \frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}. \end{aligned}$$

6.7 One-sample tests about μ (cont.)

One-sample likelihood ratio test statistic as a function of one-sample Student t -statistic.

$U_{\text{LR}} = -2 \ln(\lambda(\mathbf{X})) \stackrel{\mathcal{D}}{\sim} \chi_1^2$, where H_0 is rejected for high values of the ratio $\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}$. Then

$\hat{\sigma}_0^2 = \hat{\sigma}^2 + (\bar{x} - \mu_0)^2$ and then $\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2} = 1 + \frac{(\bar{x} - \mu_0)^2}{\hat{\sigma}^2}$. We now that

$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{n}{n-1} \hat{\sigma}^2$, then $\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}$ is increasing function of $|t_W|$ and then

$$\begin{aligned} u_{\text{LR}} &= -2 \ln(\lambda(\mathbf{x})) = n \ln \frac{\hat{\sigma}_0^2}{\hat{\sigma}^2} = n \ln \left(1 + \frac{n \sum_{i=1}^n (x_i - \mu_0)^2}{s_n^2} \right) = n \ln \left(1 + \frac{u_W}{n} \right) \\ &= n \ln \left(1 + \frac{t_W^2}{n-1} \right), \end{aligned}$$

where $u_W = n \frac{\sum_{i=1}^n (x_i - \mu_0)^2}{s_n^2} = n \frac{\sum_{i=1}^n (x_i - \mu_0)^2}{\frac{n}{n-1} s^2} = \frac{n}{n-1} t_W^2$.

H_{02} vs H_{12} . Then $u_{\text{LR}} = 0$ if $t_W \leq 0$ and $u_{\text{LR}} = \frac{n}{2} \ln(1 + \frac{t_W^2}{n-1})$ for $t_W > 0$.

H_{03} vs H_{13} . Then $u_{\text{LR}} = 0$ if $t_W \geq 0$ and $u_{\text{LR}} = \frac{n}{2} \ln(1 + \frac{t_W^2}{n-1})$ for $t_W < 0$.

6.8 One-sample tests about μ (cont.)

One-sample Score statistic as a function of one-sample Student t -statistic.

We know that $U_S = (\mathbf{a}^T S(\hat{\boldsymbol{\theta}}_0))^T \left(\mathbf{a}^T (\mathcal{J}(\hat{\boldsymbol{\theta}}_0))^{-1} \mathbf{a} \right) \mathbf{a}^T S(\hat{\boldsymbol{\theta}}_0)$, where $\mathcal{J}_{11}(\hat{\boldsymbol{\theta}}_0) = n/\hat{\sigma}_0^2$ and $S_1(\hat{\boldsymbol{\theta}}_0) = S(\mu_0) = \frac{1}{\hat{\sigma}_0^2} \sum_{i=1}^n (x_i - \mu_0)$. Then

$$\begin{aligned} u_S &= \left(\frac{1}{\hat{\sigma}_0^2} \sum_{i=1}^n (x_i - \mu_0) \right)^2 \frac{\hat{\sigma}_0^2}{n} = \left(\frac{1}{\hat{\sigma}_0^2} \left(\sum_{i=1}^n x_i - n\mu_0 \right) \right)^2 \frac{\hat{\sigma}_0^2}{n} = \frac{1}{n\hat{\sigma}_0^2} (n\bar{x} - n\mu_0)^2 \\ &= n \frac{(\bar{x} - \mu_0)^2}{\hat{\sigma}_0^2} \\ &= \frac{n(\bar{x} - \mu_0)^2}{s_n^2 + (\bar{x} - \mu_0)^2} = \frac{n(\bar{x} - \mu_0)^2}{s_n^2 \left(1 + \frac{(\bar{x} - \mu_0)^2}{s_n^2} \right)} = \frac{u_W}{1 + \frac{(\bar{x} - \mu_0)^2}{s_n^2}} = \frac{nu_W}{n + u_W} = \frac{nt_W^2}{n - 1 + t_W^2}. \end{aligned}$$

6.9 One-sample tests about μ (cont.)

One-sample Wald statistic as a function of one-sample Student t -statistic.

We know that $U_W = (\mathbf{a}^T \hat{\boldsymbol{\theta}} - \mathbf{a}^T \hat{\boldsymbol{\theta}}_0)^T \left(\mathbf{a}^T \mathcal{J}(\hat{\boldsymbol{\theta}}) \mathbf{a} \right) (\mathbf{a}^T \hat{\boldsymbol{\theta}} - \mathbf{a}^T \hat{\boldsymbol{\theta}}_0)$ where $\mathcal{J}_{11}(\hat{\boldsymbol{\theta}}) = \mathcal{J}(\hat{\mu}) = n/s_n^2$. Then

$$u_W = (\bar{x} - \mu_0)^T \frac{n}{\hat{\sigma}^2} (\bar{x} - \mu_0) = n \frac{(\bar{x} - \mu_0)^2}{s_n^2} = \frac{n}{n-1} t_W^2.$$

6.10 One-sample tests about μ (cont.)

Confidence intervals:

- **Wald** $100 \times (1 - \alpha)\%$ **empirical confidence interval for μ** is defined as

$$\mathcal{CS}_{1-\alpha}^{(W)} = \{\mu_0 : U_W(\mu) < \chi_1^2(\alpha)\}.$$

- **Likelihood** $100 \times (1 - \alpha)\%$ **empirical confidence interval for μ** is defined as

$$\mathcal{CS}_{1-\alpha}^{(LR)} = \{\mu_0 : U_{LR}(\mu) < \chi_1^2(\alpha)\}.$$

- **Score** $100 \times (1 - \alpha)\%$ **empirical confidence interval for μ** is defined as

$$\mathcal{CS}_{1-\alpha}^{(S)} = \{\mu_0 : U_S(\mu) < \chi_1^2(\alpha)\}.$$

6.11 One-sample tests about μ (cont.)

Knowing the inequality

$$\frac{x}{1+x} < \ln(1+x) < x, \text{ where } x > -1,$$

and $x = T_W^2/(n-1)$, the three test statistics can be ordered

$$U_S < U_{LR} < U_W.$$

Then

$$U_S = \frac{nT_W^2}{n-1 + T_W^2} \frac{\frac{1}{n-1}}{\frac{1}{n-1}} = \frac{n \frac{T_W^2}{n-1}}{1 + \frac{T_W^2}{n-1}} = \frac{nx}{1+x},$$

$$U_{LR} = n \ln \left(1 + \frac{T_W^2}{n-1} \right) = n \ln(1+x),$$

$$U_W = \frac{n}{n-1} T_W^2 = nx.$$

Note: If H_0 is true, $T_W = \frac{\bar{X} - \mu_0}{S} \sqrt{n} \stackrel{\mathcal{D}}{\sim} t_{n-1}$.

6.12 One-sample tests about μ (cont.)

If H_0 is true, then **the probabilities of Type I error** are

$$\alpha_W = \Pr(U_W \geq \chi_1^2(\alpha)) = \Pr\left(F_{1,n-1} \geq \frac{n-1}{n} \chi_1^2(\alpha)\right),$$

$$\alpha_{LR} = \Pr(U_{LR} \geq \chi_1^2(\alpha)) = \Pr\left(F_{1,n-1} \geq (n-1) \left[\exp\left(\frac{\chi_1^2(\alpha)}{n}\right) - 1\right]\right)$$

and

$$\alpha_S = \Pr(U_S \geq \chi_1^2(\alpha)) = \Pr\left(F_{1,n-1} \geq \frac{n-1}{n - \chi_1^2(\alpha)} \chi_1^2(\alpha)\right).$$

Note: U_W is often liberal, U_{LR} slightly liberal a U_S is usually neither liberal nor conservative. Limiting (asymptotical) distribution, i.e., for sufficiently large n , for all test statistics is χ_1^2 distribution.

6.13 One-sample tests about μ (cont.)

Let $X \sim N(\mu, \sigma^2)$, where σ^2 is not known. We would like to test the hypotheses:

- ① $H_{01} : \mu = \mu_0$ vs $H_{11} : \mu \neq \mu_0$,
- ② $H_{02} : \mu \leq \mu_0$ vs $H_{12} : \mu > \mu_0$,
- ③ $H_{03} : \mu \geq \mu_0$ vs $H_{13} : \mu < \mu_0$.

Under H_0

$$T_W = \frac{\bar{X} - \mu_0}{S} \sqrt{n} \stackrel{\mathcal{D}}{\sim} t_{n-1},$$

where $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ and the distribution t_{df} , where $df = n - 1$, is called **central t -distribution** with $n - 1$ degrees of freedom. T_W is called **one-sample Student test statistic** (or **one-sample t -statistic**) and test **one-sample Student t -test about μ** .

6.14 One-sample tests about μ (cont.)

If H_0 is not true (under the alternative H_1), this situation leads to **non-central t -distribution** with df degrees of freedom and non-centrality parameter λ , $t_{df,\lambda}$, where

$$T_{W,\lambda} = \frac{Z_W + \lambda}{\sqrt{V/df}} = \frac{\sqrt{n}(\bar{X} - \mu)/\sigma + \sqrt{n}(\mu - \mu_0)/\sigma}{S/\sigma} \stackrel{\mathcal{D}}{\sim} t_{df,\lambda},$$

$df = n - 1$, $Z_W = \frac{\bar{X} - \mu}{\sigma} \sqrt{n} \stackrel{\mathcal{D}}{\sim} N(0, 1)$, non-centrality parameter $\lambda = (\delta/\sigma) \sqrt{n}$,

$\delta = \mu - \mu_0$ is **minimal detected distance** between μ and μ_0 , $V = \frac{(n-1)S^2}{\sigma^2} \stackrel{\mathcal{D}}{\sim} \chi_{df}^2$ and is independent of Z_W . Let cumulative distribution function of $T_{W,\lambda}$ be

$G_{df,\lambda}(t) = \Pr(T_{W,\lambda} \leq t)$. If $\lambda = 0$, then non-central t -distribution is equivalent to the central (Student) t -distribution.

6.15 One-sample tests about μ (cont.)

If H_{02} is not true (under the alternative H_{12}), then the power function

$$1 - \beta(\mu, \sigma) = \Pr_{\mu, \sigma} \left(\frac{\bar{X} - \mu_0}{S} \sqrt{n} \geq t_{n-1}(\alpha) \right) = 1 - G_{n-1,\lambda}(t_{n-1}(\alpha)),$$

where (μ, σ) means that the probability is calculated under the condition $X \sim N(\mu, \sigma^2)$, $\mu > \mu_0$.

If H_{03} is not true (under the alternative H_{13}), then the power function

$$1 - \beta(\mu, \sigma) = \Pr_{\mu, \sigma} \left(\frac{\bar{X} - \mu_0}{S} \sqrt{n} \leq -t_{n-1}(\alpha) \right) = G_{n-1,\lambda}(t_{n-1}(\alpha)),$$

where (μ, σ) means that the probability is calculated under the condition $X \sim N(\mu, \sigma^2)$, $\mu < \mu_0$.

If H_{01} is not true (under the alternative H_{11}), then the power function

$$\begin{aligned} 1 - \beta(\mu, \sigma) &= \Pr_{\mu, \sigma} \left(\frac{\bar{X} - \mu_0}{S} \sqrt{n} \leq -t_{n-1}(\alpha/2) \vee \frac{\bar{X} - \mu_0}{S} \sqrt{n} \geq t_{n-1}(\alpha/2) \right) \\ &= 1 - G_{n-1,\lambda}(t_{n-1}(\alpha/2)) + G_{n-1,\lambda}(-t_{n-1}(\alpha/2)), \end{aligned}$$

where (μ, σ) means that the probability is calculated under the condition $X \sim N(\mu, \sigma^2)$, $\mu \neq \mu_0$.

6.16 One-sample tests about μ (cont.)

Critical regions and power functions related to T_W

H_0	H_1	\mathcal{W}	$1 - \beta(\mu)$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\mathcal{W}_1 = \{T_W; T_W \geq t_{n-1}(\alpha/2)\}$	$\Pr(F_{1,n-1,\lambda^2} \geq F_{1,n-1}(\alpha))$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\mathcal{W}_2 = \{T_W; T_W \geq t_{n-1}(\alpha)\}$	$\Pr(t_{n-1,\lambda} \geq t_{n-1}(\alpha))$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\mathcal{W}_3 = \{T_W; T_W \leq -t_{n-1}(\alpha)\}$	$\Pr(t_{n-1,\lambda} \leq -t_{n-1}(\alpha))$

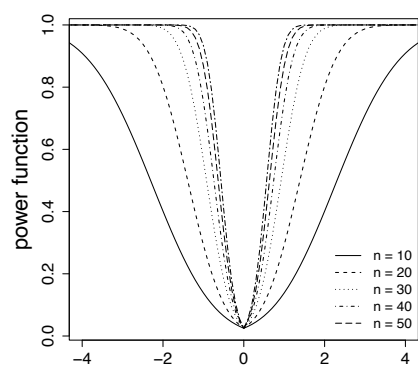
Note: $t_{n-1}^2(\alpha/2) \approx F_{1,n-1}(\alpha)$ and $t_{n-1,\lambda}^2(\alpha/2) \approx F_{1,n-1,\lambda^2}(\alpha)$.

Note: Noncentrality parameter $\lambda = (\delta/\sigma) \sqrt{n}$, $\delta = \mu - \mu_0$.

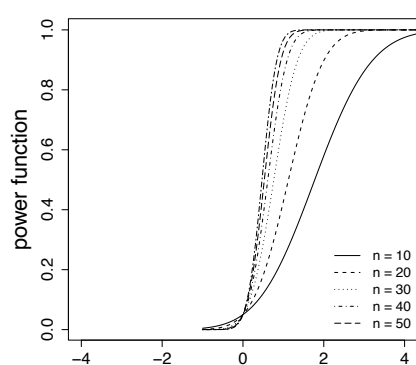
$$\text{p-value} = \begin{cases} 2\Pr(T_W \geq |t_W| | H_{01}), & \text{if } H_{11} : \mu \neq \mu_0 \\ \Pr(T_W \geq t_W | H_{02}), & \text{if } H_{12} : \mu > \mu_0 \\ \Pr(T_W \leq t_W | H_{03}), & \text{if } H_{13} : \mu < \mu_0 \end{cases}$$

H_0	H_1	$(\hat{\mu}_L, \hat{\mu}_U)$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\mathcal{CS}_{1-\alpha} = \left\{ \mu_0 : \mu_0 \in \left(\bar{x} - t_{n-1}(\alpha/2) \frac{s}{\sqrt{n}}, \bar{x} + t_{n-1}(\alpha/2) \frac{s}{\sqrt{n}} \right) \right\}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\mathcal{CS}_{1-\alpha} = \left\{ \mu_0 : \mu_0 \in \left(\bar{x} - t_{n-1}(\alpha) \frac{s}{\sqrt{n}}, \infty \right) \right\}$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\mathcal{CS}_{1-\alpha} = \left\{ \mu_0 : \mu_0 \in \left(-\infty, \bar{x} + t_{n-1}(\alpha) \frac{s}{\sqrt{n}} \right) \right\}$

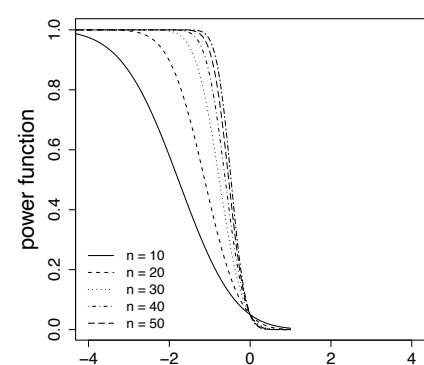
6.17 One-sample tests about μ (cont.)



$H_{11} : \mu - \mu_0 \neq 0, \mu_0 = 0, \sigma^2 = 4, \alpha = 0.05$



$H_{12} : \mu - \mu_0 > 0, \mu_0 = 0, \sigma^2 = 4, \alpha = 0.05$



$H_{13} : \mu - \mu_0 < 0, \mu_0 = 0, \sigma^2 = 4, \alpha = 0.05$

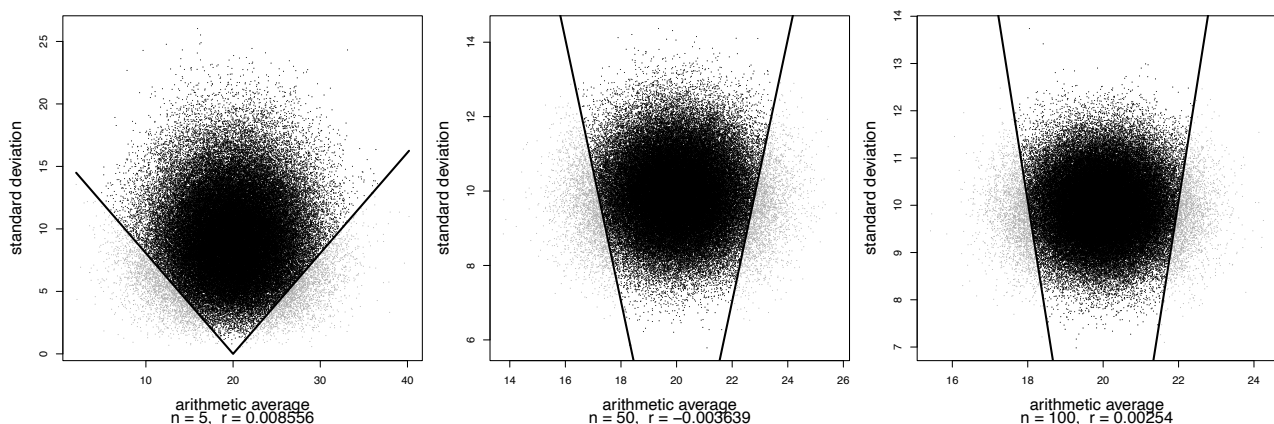
6.18 One-sample Student t -test about mean μ in R

Example (independence of μ and σ^2 , coverage probability)

Let $X \sim N(\mu, \sigma^2)$, where $\mu = 20$ and $\sigma^2 = 100$. Calculate Pearson correlation coefficient $r_{\bar{X}, S}$ based on the simulation study. Draw the points (\bar{x}_m, s_m) as scatter-plot (grey color), where $m = 1, 2, \dots, M$, $M = 100000$. Add the points

$t_{W,m} = \left| \frac{\bar{x}_m - \mu}{s_m} \sqrt{n} \right| < t_{n-1}(\alpha/2)$ (black color), and a boundary, to define the points (\bar{x}_m, s_m) , where $t_{W,m} = t_{n-1}(\alpha/2)$. Calculate coverage probability 95% of CI for μ as a ration $\sum_m I(t_{W,m} < t_{n-1}(\alpha/2)) / M$. Use (a) $n = 5$, (b) $n = 50$ and (c) $n = 100$.

6.19 One-sample Student t -test about mean μ in R

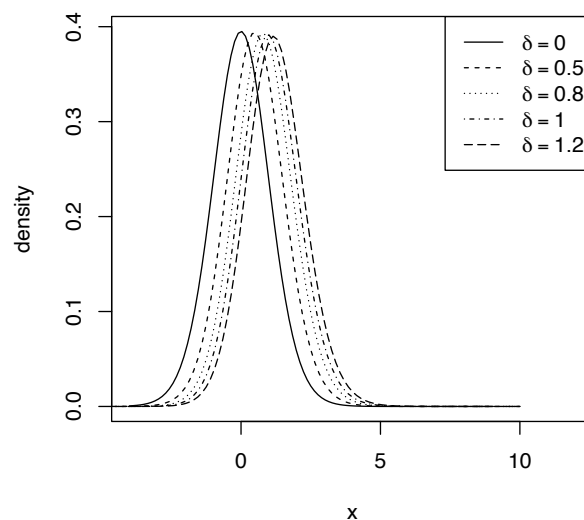


Scatter-plot of \bar{x}_m and s_m , $m = 1, 2, \dots, M$, $M = 100000$ for $n = 5$ (left), $n = 50$ (middle) and $n = 100$ (right). Lines indicate $t_{W,m} = \left| \frac{\bar{x}_m - \mu}{s_m} \sqrt{n} \right| < t_{n-1}(\alpha/2)$, where the **coverage probability** is approximately equal to the nominal level $1 - \alpha = 0.95$.

6.20 One-sample Student t -test about mean μ in R

Example (noncentral t -distribution)

Draw the densities of one central and four non-central t -distributions $t_{n-1,\lambda}$ ($\delta = \mu - \mu_0$ and $\lambda = \delta/(\sigma/\sqrt{n})$) to one figure and distinguish them by different color or line type. Use $\mu_0 = 0$, $\delta = 0, 0.5, 0.8, 1$ and 1.2 , $\sigma = 1.4$ and $n = 26$.



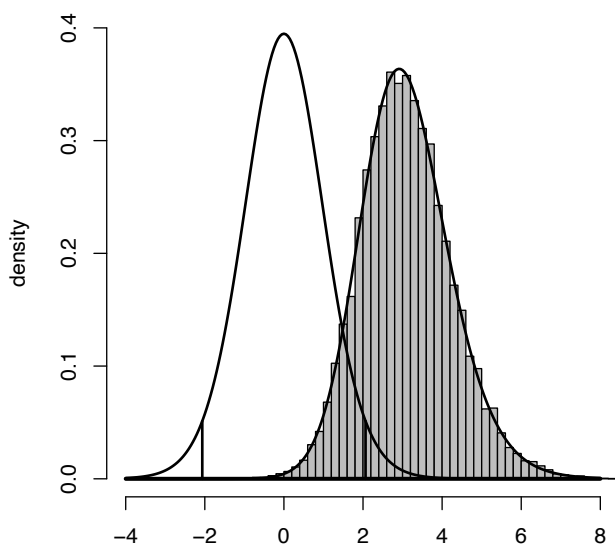
6.21 One-sample Student t -test about mean μ in R

Example (Simulated density of $t_{n-1,\lambda}$)

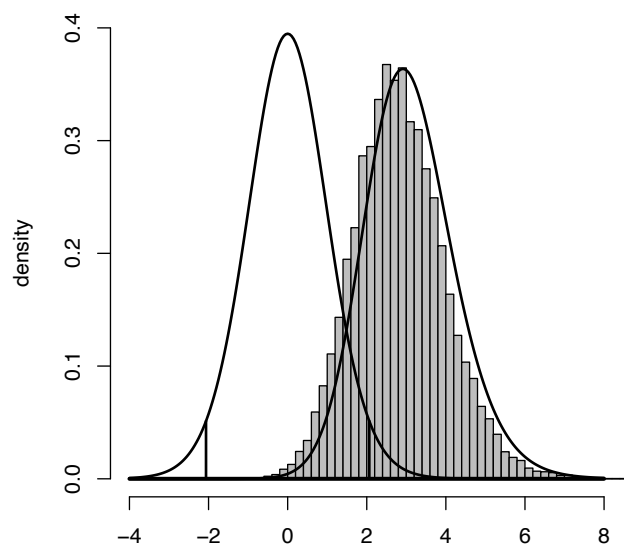
Use R to simulate the density of $t_{n-1,\lambda}$, $n = 25$, of test statistics $t_{W,\lambda}^{(m)} = \frac{\bar{x}_m - \mu_0}{s_m} \sqrt{n}$, $\lambda = 3$, $m = 1, 2, \dots, M$, with $M = 20000$ repetitions. The distribution of $t_{n-1,\lambda}$ is non-central t -distribution with $n - 1$ degrees of freedom and non-centrality parameter λ . Based on this distribution, calculate power of one-sample Student t -test for $\mu_0 = 2.5$ and $\mu_1 = 4$. Draw density of central and non-central t -distribution superimposed by histogram of simulated data. Use the following assumptions

- $X \sim N(4, 2.5^2)$ and
- $X \sim [pN(4, 2.5^2) + (1 - p)N(4, 4.5^2)]$, where $p = 0.9$.

6.22 One-sample Student t -test about mean μ in R



central and simulated non-central t -distribution
 $1 - \beta(\mu_1) = 0.8204$



central and simulated non-central t -distribution
 $1 - \beta(\mu_1) = 0.7447$

Density of central (left) and non-central (right) t -distribution superimposed by histogram of simulated data from $X \sim N(4, 2.5^2)$ and $X \sim [0.9N(4, 2.5^2) + 0.1N(4, 4.5^2)]$, resp.

6.23 One-sample Student t -test about mean μ in R

Probability of empirical Type I error for MC experiment is the probability p of statistically significant test statistics between M repetitions, if H_0 is true. Then $SD(p) = \sqrt{p(1-p)/M} \leq 0.5/\sqrt{M}$.

Example (Probability of empirical Type I error)

Let $X \sim N(\mu, \sigma^2)$, where $\mu = 500$ and $\sigma^2 = 100$. Test $H_0 : \mu \leq 500$ vs $H_1 : \mu > 500$, where $\alpha = 0.05$, σ is not known. Use R to simulate empirical $\Pr(\text{Type I error})$, where the number of simulations $M = 10000$ and the sample size $n = 20$ for one-sample Student t -test about mean μ . Use the function `t.test(x, alternative = "greater", mu = mu0)` to calculate p-value and standard deviation for each test statistic $t_m, m = 1, 2, \dots, M$ under H_0 . The p-value is relative frequency of p of rejected H_0 on $\alpha = 0.05$ between M tests, i.e., $p = \Pr(\text{Type I error}) = \frac{\sum_{m=1}^M I(H_0 \text{ reject})}{M}$.

6.24 One-sample Student t -test about mean μ in R

```

1  n <- 20
2  alfa <- .05
3  mu0 <- 500
4  sigma <- 100
5  M <- 10000
6  p_values <- numeric(M)
7  for (j in 1:M) {
8    x <- rnorm(n, mu0, sigma)
9    ttest <- t.test(x, alternative = "greater", mu = mu0)
10   p_values[j] <- ttest$p.value
11 }
12 p.hat <- mean(p_values < alfa)
13 se.hat <- sqrt(p.hat * (1 - p.hat) / M)
14 print(c(p.hat, se.hat)) # 0.052100000 0.002222287

```

```
[1] 0.048800000 0.002154497
```

6.25 One-sample Student t -test about mean μ in R

Example (Empirical power function of one-sample Student t -test)

Use R to simulate empirical power function of one-sample Student t -test about μ . Test $H_0 : \mu = 500$ vs $H_1 : \mu \neq 500$, resp. $\mu = \mu_1 = 450, 460, \dots, 640, 650$ (two-sided alternative). Use the function `t.test(x, mu = 500)`, to calculate

$t_m = \frac{\mu_1 - \mu_0}{s_m} \sqrt{n}$, $m = 1, 2, \dots, M$, where $M = 10000$, $n = 20$, p-values related to t_m and

compare them with $\alpha = 0.05$. $1 - \hat{\beta}(\mu_1)$ for a given alternative is then divided as number of these p-values divided by M . Draw $1 - \hat{\beta}(\mu_1)$ for a given alternative and also their

standard deviations $SD[\widehat{1 - \hat{\beta}(\mu_1)}] = \sqrt{\frac{(1 - \hat{\beta}(\mu_1))\hat{\beta}(\mu_1)}{M}}$ as a vertical error bar

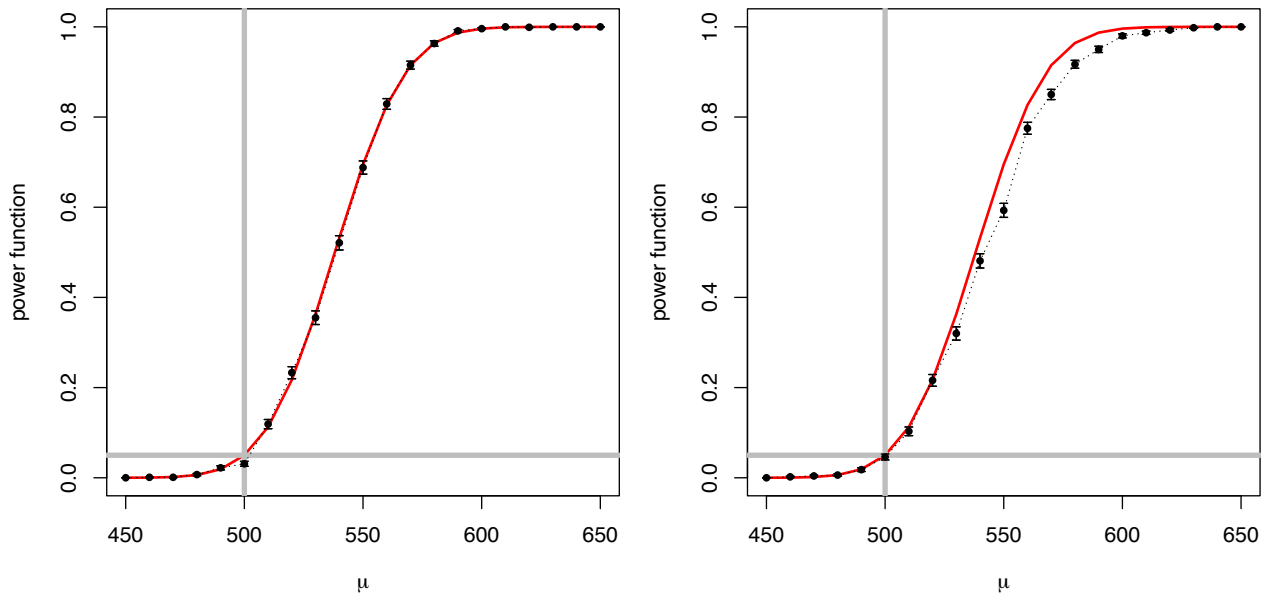
$1 - \hat{\beta}(\mu_1) \pm SD[\widehat{1 - \hat{\beta}(\mu_1)}]$. Draw also theoretical power function

$1 - \beta(\mu_1)$, $\mu_1 \in \langle 450, 650 \rangle$ (use the function `power.t.test()`). Use the following assumptions

A. $X \sim N(\mu_1, 100^2)$ and

B. $X \sim [pN(\mu_1, 100^2) + (1 - p)N(\mu_1, 200^2)]$, where $p = 0.9$.

6.26 One-sample Student t -test about mean μ in R



Empirical power function (grey smooth curve as mean of values characterised by boxplots) and **theoretical power function** (red smooth curve) of **one-sample Student t -statistics** T_W , left – simulated data from $X \sim N(500, 100^2)$. right – simulated data from $X \sim [0.9N(500, 100^2) + 0.1N(500, 200^2)]$.

6.27 One-sample tests about μ (cont.)

Example (one-sample tests for testing hypotheses about mean μ)

Data: one-sample-mean-skull-mf.txt and variable X skull length skull.L in millimeters (mm) of ancient Egyptian male population.

Expectation: normal distribution of this variable, i.e., $X \sim N(\mu, \sigma^2)$.

- Test null hypothesis (against general alternative) about equality of mean μ of skull length of this population with mean equal to $\mu_0 = 177.568$ mm of modern Egyptian male population on a significance level $\alpha = 0.05$.
- Calculate $100 \times (1 - \alpha)\%$ empirical confidence interval for the mean, where the coverage probability (confidence coefficient) is equal to $1 - \alpha = 0.95$.

Note: Use (1) Wald test statistic T_W , (2) Likelihood ratio test statistic U_{LR} and equivalent confidence intervals. **Results of Wald test compare with the results using `t.test()` function.**

6.28 One-sample tests about μ (cont.)

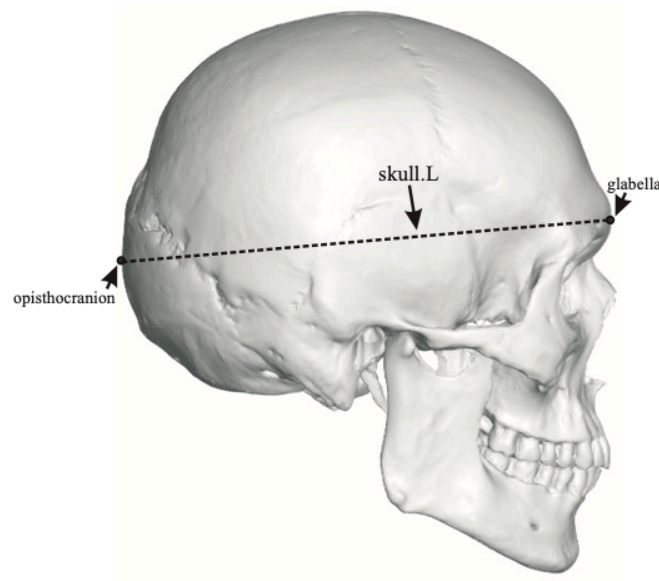


Figure 2: Skull length measured in millimeters (mm)

6.29 One-sample tests about μ (cont.)

Mathematical formulation:

- **Null hypothesis** – $H_0 : \mu = \mu_0$
- **Alternative hypothesis** – $H_1 : \mu \neq \mu_0$,

where $\mu_0 = 177.568$

Verbal formulation:

- **Null hypothesis** – mean μ of skull length of ancient Egyptian male population is equal to the mean $\mu_0 = 177.568$ mm of modern Egyptian male population.
- **Alternative hypothesis** – mean μ of skull length of ancient Egyptian male population is not equal to the mean $\mu_0 = 177.568$ mm of modern Egyptian male population.

6.30 One-sample tests about μ (cont.)

The elements of testing procedure:

- **Realisation of test statistic** – Mean μ and variance σ^2 are not known and therefore we need to estimate them. The estimates are arithmetic average (maximal likelihood estimate) $\bar{x} \doteq 182.037$ and variance (unbiased estimate) $s^2 \doteq 40.582$ (standard deviation $s \doteq 6.370$). Then
 - Realisation of **Wald test statistic** $t_W = \frac{\bar{x} - \mu_0}{s} \sqrt{n} = \frac{182.037 - 177.568}{6.370} \sqrt{217} \doteq 10.334$
 - Realisation of **likelihood ratio test statistic** $u_{LR} = n \ln \left(1 + \frac{t_W^2}{n-1} \right) \doteq 87.172$
- **Rejection region** –
 - **Wald test**
critical values $t_{n-1}(1 - \alpha/2) = t_{216}(1 - 0.025) \doteq -1.971$ and $t_{n-1}(\alpha/2) = t_{216}(0.025) \doteq 1.971$
rejection region $\mathcal{W}_1 = (t_{\min}, t_{n-1}(1 - \alpha/2)) \cup (t_{n-1}(\alpha/2), t_{\max}) = (-\infty, -1.971) \cup (1.971, \infty)$
 - **likelihood ratio test**
critical value $\chi_1^2(\alpha) \doteq 3.841$
rejection region $\mathcal{W} = (\chi_1^2(\alpha), t_{\max}) \doteq (3.841, \infty)$

6.31 One-sample tests about μ (cont.)

The elements of testing procedure (cont.):

- **95% empirical confidence interval for the mean**
 - **Wald 95% empirical confidence interval for μ**
 $(\hat{\mu}_L, \hat{\mu}_U) = (\bar{x} - t_{n-1}(1 - \alpha/2) \frac{s}{\sqrt{n}}, \bar{x} + t_{n-1}(\alpha/2) \frac{s}{\sqrt{n}}) \doteq (182.037 - 1.971 \frac{6.370}{\sqrt{30}}, 182.037 + 1.971 \frac{6.370}{\sqrt{30}}) = (181.185, 182.890)$
 - **likelihood 95% empirical confidence interval for μ**
 $\mathcal{CS}_{0.95} = \{\mu_0 : u_{LR}(\mu_0) < \chi_1^2(\alpha)\} \doteq (181.188, 182.886)$
- **p-values** –
 - **Wald test** p-value $\doteq 2\Pr(T_W \geq |10.334| | H_0) < 0.0001$
 - **likelihood ratio test** p-value $\doteq \Pr(U_{LR} \geq 87.172 | H_0) < 0.0001$
- **Statistical conclusion** – H_0 is rejected on a significance level $\alpha = 0.05$, because (1) tests statistic belong to the critical region, (2) $\mu_0 = 177.568$ does not belong to the confidence interval, and (3) p-value is smaller than 0.05.
- **Verbal conclusion** – We are rejecting null hypothesis that mean μ of skull length of ancient Egyptian male population is equal to the mean $\mu_0 = 177.568$ mm of modern Egyptian male population.

6.32 One-sample tests about μ (cont.)

```

1 wd <- "~/Documents/TEACHING/LECTURES/M8986-SI-II/DATA/"
2 fname <- paste(wd, "one-sample-mean-skull-mf.txt", sep = "")
3 DATA <- read.table(fname, header = TRUE)
4 skull.L <- DATA$skull.L[DATA$sex == "m"]
5
6 x <- na.omit(skull.L)
7 n <- length(x) ## 217
8 mu_hat <- mean(x) ## 182.0369
9 var_hat <- var(x) ## 40.58197
10 sd_hat <- sd(x) ## 6.370398
11 mu0 <- 177.568
12 tW_obs <- (mu_hat - mu0)/sd_hat*sqrt(n) ## 10.33382
13 sigma_sq_hat <- (n - 1)/n*var(x) ## 40.58197
14 sigma0_sq_hat <- sum((x - mu0)^2)/n ## 60.36572
15 uLR_obs <- n*log(sigma0_sq_hat/sigma_sq_hat) ## 87.17249
16 t_critval_d <- qt(0.025, df = n - 1) ## -1.971007
17 t_critval_h <- qt(0.975, df = n - 1) ## 1.971007

```

6.33 One-sample tests about μ (cont.)

```

1 lr_crit_cal <- qchisq(0.95, df = 1) ## 3.841459
2 CI_W <- mu_hat + c(-1, 1)*t_critval_h*sd_hat/sqrt(n) ## 181.1845 182.889
3 min_mu <- mu_hat - 1*sd(x)
4 max_mu <- mu_hat + 1*sd(x)
5 mu_0i <- seq(min_mu, max_mu, by = 0.0001)
6 tW_obs_i <- (mu_hat - mu_0i)/sd(x)*sqrt(n)
7 uW_obs_i <- tW_obs_i^2
8 uLR_i <- n*log(1 + uW_obs_i/(n - 1))
9 CI_LR <- range(mu_0i[which(uLR_i < lr_crit_cal)]) ## 181.1876 182.8862
10 p_value_W <- 2*(1 - pt(abs(tW_obs), df = n - 1)) ## < 0.0001
11 p_value_LR <- 1 - pchisq(uLR_obs, df = 1) ## < 0.0001
12
13 W_test <- t.test(x, mu = mu0)
14 # W_test$estimate ## 182.0369
15 # W_test$conf.int ## 181.1845 182.8892
16 # W_test$stat ## 10.33382
17 # W_test$p.val ## 1.344108e-20

```

6.34 One-sample tests about μ (cont.)

```

1  # Likelihood CI for mu using Brent method
2  "ULR" <- function(mu_0,mu_hat) {
3    tW_obs_i <- (mu_hat - mu_0)/sd(x)*sqrt(n)
4    uW_obs_i <- tW_obs_i^2
5    uLR_i <- n*log(1 + uW_obs_i/(n - 1))
6    uLR_i - qchisq(0.95,1)
7  }
8  likeCI_mu_brent <- c(uniroot(ULR,c(180,mu_hat),mu_hat = mu_hat)$root,
9                      uniroot(ULR,c(mu_hat,184),mu_hat = mu_hat)$root)
10 ## 181.1875 182.8863

```

6.35 One-sample Student t -test about mean μ in , function `t.test()`

Arguments (inputs):

- ❶ data vector \mathbf{x} – `x`
- ❷ alternative – default `alternative = "two.sided"`, other choices `"greater"`, `"less"`
- ❸ mean under H_0 , i.e. μ_0 – `mu`
- ❹ **one-sample approach** – `paired = FALSE`
- ❺ coverage probability $1 - \alpha$ – `conf.level`, default 0.95

Outputs:

- ❶ test statistic t_W – `statistic`
- ❷ degrees of freedom (df) – `parameter`
- ❸ p-value – `p.value`
- ❹ alternative hypothesis – `alternative hypothesis`
- ❺ empirical confidence interval for μ – `conf.int`
- ❻ point estimate (arithmetic average \bar{x}) – `sample estimates`
- ❼ name of the test – `method`