Dataanalyse - Hypotesetest - Kursusgang 3

Ege Rubak - rubak@math.aau.dk
http://www.math.aau.dk/∼rubak/teaching/2010/nano4

19. februar 2010

Confidence interval for μ with known σ

Sample with $(X_1, \ldots, X_n)$ independent, $X_i \sim N(\mu, \sigma^2)$.

- We have: $\overline{X} \sim N(\mu, \sigma^2/n)$ and

$$Z = \frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1).$$

- Therefore:

$$P\left(z_{\alpha/2} \leq Z \leq z_{1-\alpha/2}\right) = 1 - \alpha,$$

  where $z_{\alpha/2}$ is the $\alpha/2$ quantile in $N(0,1)$. I.e. if α is 0.05 as usual we look up the 2.5% quantile $z_{0.025} = -1.96$ and the 97.5% quantile $z_{0.975} = 1.96$ in the standard normal distribution $N(0,1)$.
- I.e. the confidence interval is:

$$[\overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}; \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]$$

Confidence interval for μ with known σ

Sample with $(X_1, \dots, X_n)$ independent, $X_i \sim N(\mu, \sigma^2)$.

- We have: $\overline{X} \sim N(\mu, \sigma^2/n)$ and

$$Z = \frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1).$$

- Therefore:

$$P\left(z_{\alpha/2} \leq Z \leq z_{1-\alpha/2}\right) = 1 - \alpha,$$

  where $z_{\alpha/2}$ is the $\alpha/2$ quantile in $N(0,1)$. I.e. if α is 0.05 as usual we look up the 2.5% quantile $z_{0.025} = -1.96$ and the 97.5% quantile $z_{0.975} = 1.96$ in the standard normal distribution $N(0,1)$.

- I.e. the confidence interval is:

$$[\overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}; \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]$$

Confidence interval for μ with known σ

Sample with $(X_1, \ldots, X_n)$ independent, $X_i \sim N(\mu, \sigma^2)$.

- We have: $\overline{X} \sim N(\mu, \sigma^2/n)$ and

$$Z = \frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1).$$

- Therefore:

$$P\left(z_{\alpha/2} \leq Z \leq z_{1-\alpha/2}\right) = 1 - \alpha,$$

  where $z_{\alpha/2}$ is the $\alpha/2$ quantile in $N(0,1)$. I.e. if α is 0.05 as usual we look up the 2.5% quantile $z_{0.025} = -1.96$ and the 97.5% quantile $z_{0.975} = 1.96$ in the standard normal distribution $N(0,1)$.
- I.e. the confidence interval is:

$$[\overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}; \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]$$

Hypothesis test for $\mu = \mu_0$ with known σ

- Suppose we have a hypothesis H_0: $\mu = \mu_0$ (μ_0 is a known number). Assuming H_0 is true we can calculate

$$Z = \frac{\overline{X} - \mu_0}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

for a given data set. This is called the Z-statistic, and if $Z \notin [z_{\alpha/2}, z_{1-\alpha/2}]$ we reject H_0.

- This is the same as checking μ_0 is in the confidence interval

$$[\overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}; \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]$$

- Sometimes called a z-test (e.g. in MATLAB).

Hypothesis test for $\mu = \mu_0$ with known σ

- Suppose we have a hypothesis H_0: $\mu = \mu_0$ (μ_0 is a known number). Assuming H_0 is true we can calculate

$$Z = \frac{\overline{X} - \mu_0}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

for a given data set. This is called the Z-statistic, and if $Z \notin [z_{\alpha/2}, z_{1-\alpha/2}]$ we reject H_0.

- This is the same as checking μ_0 is in the confidence interval

$$[\overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}; \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]$$

- Sometimes called a z-test (e.g. in MATLAB).

Hypothesis test for $\mu = \mu_0$ with known σ

- Suppose we have a hypothesis H_0: $\mu = \mu_0$ (μ_0 is a known number). Assuming H_0 is true we can calculate

$$Z = \frac{\overline{X} - \mu_0}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

  for a given data set. This is called the Z-statistic, and if $Z \notin [z_{\alpha/2}, z_{1-\alpha/2}]$ we reject H_0.
- This is the same as checking μ_0 is in the confidence interval

$$[\overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}; \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]$$

- Sometimes called a z-test (e.g. in MATLAB).

When σ is unknown

- Same same, but different!
- Since σ^2 is unknown we use the estimator, $S^2 \sim \frac{\sigma^2}{n-1}\chi^2(n-1)$, and we know

$$T = \frac{\overline{X} - \mu}{\sqrt{S^2/n}} \sim t(n-1). \quad (1)$$

- The confidence interval is:

$$[\overline{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}; \overline{x} + t_{1-\alpha/2}\frac{s}{\sqrt{n}}] \quad (2)$$

  where $t_{\alpha/2}$ is the $\alpha/2$ quantile in $t(n-1)$.
- To test a hypothesis H_0: $\mu = \mu_0$ we insert μ_0 in (1) and reject H_0 if the T-statistic is outside the interval $[t_{\alpha/2}, t_{1-\alpha/2}]$. This is the same as checking μ_0 is in the confidence interval (2).
- Sometimes called a t-test.

When σ is unknown

- Same same, but different!
- Since σ^2 is unknown we use the estimator, $S^2 \sim \frac{\sigma^2}{n-1}\chi^2(n-1)$, and we know
$$T = \frac{\overline{X} - \mu}{\sqrt{S^2/n}} \sim t(n-1). \quad (1)$$
- The confidence interval is:
$$[\overline{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}; \overline{x} + t_{1-\alpha/2}\frac{s}{\sqrt{n}}] \quad (2)$$
where $t_{\alpha/2}$ is the $\alpha/2$ quantile in $t(n-1)$.
- To test a hypothesis H_0: $\mu = \mu_0$ we insert μ_0 in (1) and reject H_0 if the T-statistic is outside the interval $[t_{\alpha/2}, t_{1-\alpha/2}]$. This is the same as checking μ_0 is in the confidence interval (2).
- Sometimes called a t-test.

When σ is unknown

- Same same, but different!
- Since σ^2 is unknown we use the estimator, $S^2 \sim \frac{\sigma^2}{n-1}\chi^2(n-1)$, and we know

$$T = \frac{\overline{X} - \mu}{\sqrt{S^2/n}} \sim t(n-1). \tag{1}$$

- The confidence interval is:

$$[\overline{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}; \overline{x} + t_{1-\alpha/2}\frac{s}{\sqrt{n}}] \tag{2}$$

where $t_{\alpha/2}$ is the $\alpha/2$ quantile in $t(n-1)$.

- To test a hypothesis H_0: $\mu = \mu_0$ we insert μ_0 in (1) and reject H_0 if the T-statistic is outside the interval $[t_{\alpha/2}, t_{1-\alpha/2}]$. This is the same as checking μ_0 is in the confidence interval (2).
- Sometimes called a t-test.

When σ is unknown

- Same same, but different!
- Since σ^2 is unknown we use the estimator, $S^2 \sim \frac{\sigma^2}{n-1}\chi^2(n-1)$, and we know

$$T = \frac{\overline{X} - \mu}{\sqrt{S^2/n}} \sim t(n-1). \tag{1}$$

- The confidence interval is:

$$[\overline{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}; \overline{x} + t_{1-\alpha/2}\frac{s}{\sqrt{n}}] \tag{2}$$

where $t_{\alpha/2}$ is the $\alpha/2$ quantile in $t(n-1)$.

- To test a hypothesis H_0: $\mu = \mu_0$ we insert μ_0 in (1) and reject H_0 if the T-statistic is outside the interval $[t_{\alpha/2}, t_{1-\alpha/2}]$. This is the same as checking μ_0 is in the confidence interval (2).
- Sometimes called a t-test.

When σ is unknown

- Same same, but different!
- Since σ^2 is unknown we use the estimator, $S^2 \sim \frac{\sigma^2}{n-1}\chi^2(n-1)$, and we know

$$T = \frac{\overline{X} - \mu}{\sqrt{S^2/n}} \sim t(n-1). \tag{1}$$

- The confidence interval is:

$$[\overline{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}; \overline{x} + t_{1-\alpha/2}\frac{s}{\sqrt{n}}] \tag{2}$$

  where $t_{\alpha/2}$ is the $\alpha/2$ quantile in $t(n-1)$.
- To test a hypothesis H_0: $\mu = \mu_0$ we insert μ_0 in (1) and reject H_0 if the T-statistic is outside the interval $[t_{\alpha/2}, t_{1-\alpha/2}]$. This is the same as checking μ_0 is in the confidence interval (2).
- Sometimes called a t-test.

Example of hypothesis test

- Suppose we measure the height of 100 people

x_1	x_2	$\cdots$	x_{100}
178 cm	183 cm	$\cdots$	175 cm

  and calculate the estimate of the mean $\overline{x} = 178$ and the estimate of the variance $s^2 = 64$ (i.e. $s = 8$).
- We want to test $H_0 : \mu = 180$ cm at level of significance $\alpha = 0.05$.
- Since

$$t = \frac{178 - 180}{8/10} = -2.5, \quad t_{0.025} = -1.98, \quad t_{0.975} = 1.98$$

  we reject H_0.
- What about with level of significance $\alpha = 0.01$? In this case

$$t_{0.005} = -2.63, \quad t_{0.995} = 2.63,$$

  and we cannot reject H_0.

Example of hypothesis test

- Suppose we measure the height of 100 people

x_1	x_2	$\cdots$	x_{100}
178 cm	183 cm	$\cdots$	175 cm

  and calculate the estimate of the mean $\overline{x} = 178$ and the estimate of the variance $s^2 = 64$ (i.e. $s = 8$).
- We want to test $H_0 : \mu = 180$ cm at level of significance $\alpha = 0.05$.
- Since

$$t = \frac{178 - 180}{8/10} = -2.5, \quad t_{0.025} = -1.98, \quad t_{0.975} = 1.98$$

  we reject H_0.
- What about with level of significance $\alpha = 0.01$? In this case

$$t_{0.005} = -2.63, \quad t_{0.995} = 2.63,$$

  and we cannot reject H_0.

Example of hypothesis test

- Suppose we measure the height of 100 people

x_1	x_2	$\cdots$	x_{100}
178 cm	183 cm	$\cdots$	175 cm

  and calculate the estimate of the mean $\overline{x} = 178$ and the estimate of the variance $s^2 = 64$ (i.e. $s = 8$).
- We want to test $H_0 : \mu = 180$ cm at level of significance $\alpha = 0.05$.
- Since

$$t = \frac{178 - 180}{8/10} = -2.5, \quad t_{0.025} = -1.98, \quad t_{0.975} = 1.98$$

  we reject H_0.
- What about with level of significance $\alpha = 0.01$? In this case

$$t_{0.005} = -2.63, \quad t_{0.995} = 2.63,$$

  and we cannot reject H_0.

Example of hypothesis test

- Suppose we measure the height of 100 people

x_1	x_2	$\cdots$	x_{100}
178 cm	183 cm	$\cdots$	175 cm

  and calculate the estimate of the mean $\overline{x} = 178$ and the estimate of the variance $s^2 = 64$ (i.e. $s = 8$).
- We want to test $H_0 : \mu = 180$ cm at level of significance $\alpha = 0.05$.
- Since

$$t = \frac{178 - 180}{8/10} = -2.5, \quad t_{0.025} = -1.98, \quad t_{0.975} = 1.98$$

  we reject H_0.
- What about with level of significance $\alpha = 0.01$? In this case

$$t_{0.005} = -2.63, \quad t_{0.995} = 2.63,$$

  and we cannot reject H_0.

p-value

- We want a number telling how much data confirms the hypothesis.
- Define p-value:

$$
\begin{aligned}
p\text{-value} &= P(\text{similar dataset supports } H_0 \text{ less} \mid H_0 \text{ is true}) \\
&= P\left(\begin{matrix}\text{estimates from similar data-} \\ \text{set are more extreme}\end{matrix} \;\middle|\; H_0 \text{ is true}\right)
\end{aligned}
$$

p-value

- We want a number telling how much data confirms the hypothesis.
- Define *p*-value:

$$
\begin{aligned}
p\text{-value} &= P(\text{similar dataset supports } H_0 \text{ less} \mid H_0 \text{ is true}) \\
&= P\left(\begin{array}{l}\text{estimates from similar data-}\\ \text{set are more extreme}\end{array} \,\middle|\, H_0 \text{ is true}\right)
\end{aligned}
$$

p-value

- We want a number telling how much data confirms the hypothesis.
- Define *p*-value:

$$\begin{aligned} p\text{-value} &= P(\text{similar dataset supports } H_0 \text{ less} \mid H_0 \text{ is true}) \\ &= P\left(\begin{matrix}\text{estimates from similar data-} \\ \text{set are more extreme}\end{matrix} \,\middle|\, H_0 \text{ is true}\right) \\ &= P\left(\begin{matrix}\text{estimates from similar dataset} \\ \text{are more extreme if we assume} \\ H_0 \text{ is true}\end{matrix}\right) \end{aligned}$$

remember definition

p-value

- If $H_0 : \mu = 180$ is true:

$$T = \frac{\overline{X} - 180}{\sqrt{S^2/n}} \sim t(99)$$

- p-value for our observation $\overline{x} = 178$, $s^2 = 64$:

$$p\text{-value} = P\left(\begin{array}{l}\text{estimates from similar data-}\\ \text{set are more extreme}\end{array} \middle| H_0 \text{ is true}\right)$$

- NOTE: "more extreme" depends on the hypothesis.

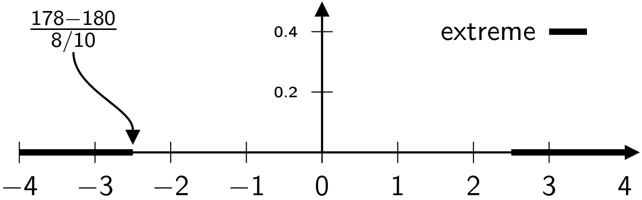
p-value

- If $H_0 : \mu = 180$ is true:

$$T = \frac{\overline{X} - 180}{\sqrt{S^2/n}} \sim t(99)$$

- *p*-value for our observation $\overline{x} = 178$, $s^2 = 64$:

$$p\text{-value} = P\left(\begin{array}{l}\text{estimates from similar data-}\\ \text{set are more extreme}\end{array} \middle| \; H_0 \text{ is true}\right)$$
$$= P(|T| \geq 2.5)$$



- NOTE: "more extreme" depends on the hypothesis.

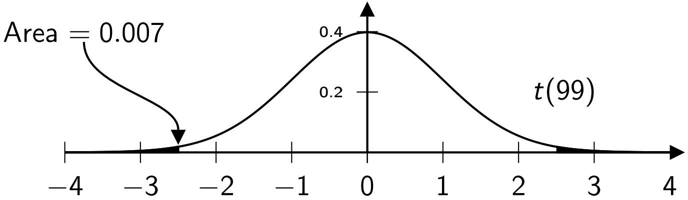
p-value

- If $H_0 : \mu = 180$ is true:

$$T = \frac{\overline{X} - 180}{\sqrt{S^2/n}} \sim t(99)$$

- p-value for our observation $\overline{x} = 178$, $s^2 = 64$:

$$p\text{-value} = P\left(\begin{array}{l}\text{estimates from similar data-}\\ \text{set are more extreme}\end{array} \middle| \; H_0 \text{ is true}\right)$$
$$= P(|T| \geq 2.5)$$



- NOTE: "more extreme" depends on the hypothesis.

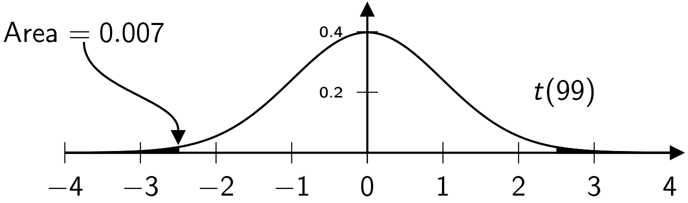
p-value

- If $H_0 : \mu = 180$ is true:

$$T = \frac{\overline{X} - 180}{\sqrt{S^2/n}} \sim t(99)$$

- *p*-value for our observation $\overline{x} = 178$, $s^2 = 64$:

$$\begin{aligned} p\text{-value} &= P\left(\begin{matrix}\text{estimates from similar data-} \\ \text{set are more extreme}\end{matrix} \;\middle|\; H_0 \text{ is true}\right) \\ &= P(|T| \geq 2.5) = 1.4\% \end{aligned}$$



- NOTE: "more extreme" depends on the hypothesis.

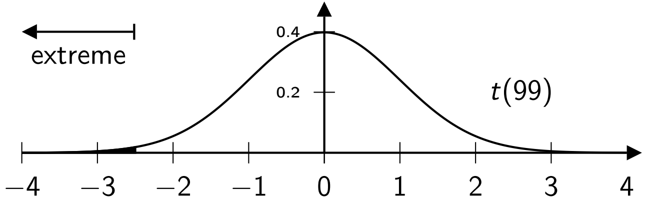
p-value

- If $H_0 : \mu \geq 180$ is true:

$$T = \frac{\overline{X} - 180}{\sqrt{S^2/n}} \sim t(99)$$

- *p*-value for our observation $\overline{x} = 178$, $s^2 = 64$:

$$p\text{-value} = P\left(\begin{array}{l}\text{estimates from similar data-}\\\text{set are more extreme}\end{array} \middle| \; H_0 \text{ is true}\right)$$
$$= 1.4\% = P(T \leq -2.5) = 0.7\%$$



- NOTE: "more extreme" depends on the hypothesis.

Fatcs about *p*-value

- Small *p*-value (≤ 0.05): We reject hypothesis.
- Large *p*-value: We cannot reject hypothesis.
- The larger the *p*-value, the more faith we have in our hypothesis.
- *p*-value (much) greater than 0.05 $\Rightarrow$ estimate is (well) inside 95 % confidence interval.

Paired samples

- Data:

$$\begin{array}{lcccc} \text{Sample 1:} & x_{1,1} & x_{1,2} & \dots & x_{1,n} \\ \text{Sample 2:} & x_{2,1} & x_{2,2} & \dots & x_{2,n} \end{array}$$

- Assumptions:
  - Observations occur in pairs, $(x_{1,i}, x_{2,i})$.
  - Each sample consists of independent, normally distributed observations, $X_{i,j} \sim N(\mu_i, \sigma_i^2)$.
- Note:
  - The two samples do not need to be independent.
  - Is often used in before-after experiments.
- Hypothesis:

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

Paired samples

- Data:

$$\begin{array}{lcccc} \text{Sample 1:} & x_{1,1} & x_{1,2} & \dots & x_{1,n} \\ \text{Sample 2:} & x_{2,1} & x_{2,2} & \dots & x_{2,n} \end{array}$$

- Assumptions:
  - Observations occur in pairs, $(x_{1,i}, x_{2,i})$.
  - Each sample consists of independent, normally distributed observations, $X_{i,j} \sim N(\mu_i, \sigma_i^2)$.
- Note:
  - The two samples do not need to be independent.
  - Is often used in before-after experiments.
- Hypothesis:

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

Paired samples

- Data:

$$\begin{array}{rcccc} \text{Sample 1}: & x_{1,1} & x_{1,2} & \dots & x_{1,n} \\ \text{Sample 2}: & x_{2,1} & x_{2,2} & \dots & x_{2,n} \\ \hline \text{Difference:} & d_1 & d_2 & \dots & d_n \end{array}$$

$d_i = x_{1,i} - x_{2,i}$

- We have

don't worry about this one

$$d_i \sim\sim N(\underbrace{\mu_1 - \mu_2}_{=\delta}, \sigma_*^2)$$

- Hypothesis:

$$\left.\begin{array}{l} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{array}\right.$$

- Use usual t-test to test if $\delta = 0$.

Paired samples

- Data:

$$\begin{array}{lcccc} \text{Sample 1}: & x_{1,1} & x_{1,2} & \dots & x_{1,n} \\ \text{Sample 2}: & x_{2,1} & x_{2,2} & \dots & x_{2,n} \\ \hline \text{Difference}: & d_1 & d_2 & \dots & d_n \end{array}$$

$d_i = x_{1,i} - x_{2,i}$

- We have

don't worry about this one

$$d_i \sim\sim N(\underbrace{\mu_1 - \mu_2}_{=\delta}, \sigma_*^2)$$

- Hypothesis:

$$\left.\begin{array}{l} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{array}\right\} \quad \Leftrightarrow \quad \left\{\begin{array}{l} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{array}\right.$$

- Use usual t-test to test if $\delta = 0$.

Paired samples

- Data:

$$\begin{array}{rcccc} \text{Sample 1}: & x_{1,1} & x_{1,2} & \dots & x_{1,n} \\ \text{Sample 2}: & x_{2,1} & x_{2,2} & \dots & x_{2,n} \\ \hline \text{Difference}: & d_1 & d_2 & \dots & d_n \end{array}$$

$d_i = x_{1,i} - x_{2,i}$

- We have

don't worry about this one

$$d_i \sim\sim N(\underbrace{\mu_1 - \mu_2}_{=\delta}, \sigma_*^2)$$

- Hypothesis:

$$\left.\begin{array}{l} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} H_0 : \delta = 0 \\ H_1 : \delta \neq 0 \end{array}\right.$$

- Use usual t-test to test if $\delta = 0$.

Paired samples

- Data:

$$\begin{array}{lcccc} \text{Sample 1}: & x_{1,1} & x_{1,2} & \dots & x_{1,n} \\ \text{Sample 2}: & x_{2,1} & x_{2,2} & \dots & x_{2,n} \\ \hline \text{Difference:} & d_1 & d_2 & \dots & d_n \end{array}$$

$d_i = x_{1,i} - x_{2,i}$

- We have

don't worry about this one

$$d_i \sim\sim N(\underbrace{\mu_1 - \mu_2}_{=\delta}, \sigma_*^2)$$

- Hypothesis:

$$\left.\begin{array}{l} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} H_0 : \delta = 0 \\ H_1 : \delta \neq 0 \end{array}\right.$$

- Use usual t-test to test if $\delta = 0$.

Unpaired samples

- Data:

$$\begin{array}{lllll} \text{Sample 1:} & x_{1,1} & x_{1,2} & \dots & x_{1,n_1} \\ \text{Sample 2:} & x_{2,1} & x_{2,2} & \dots & x_{2,n_2} \end{array}$$

- Assumptions:
  - Each sample consists of independent normally distributed observations, $X_{i,j} \sim N(\mu_i, \sigma_i^2)$.
  - The two samples are independent.
- Hypothesis:

$$\left.\begin{array}{l} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{array}\right\} \quad \Leftrightarrow \quad \left\{\begin{array}{l} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{array}\right.$$

Unpaired samples

- Data:

$$\begin{array}{lllll} \text{Sample 1:} & x_{1,1} & x_{1,2} & \dots & x_{1,n_1} \\ \text{Sample 2:} & x_{2,1} & x_{2,2} & \dots & x_{2,n_2} \end{array}$$

- Assumptions:
  - Each sample consists of independent normally distributed observations, $X_{i,j} \sim N(\mu_i, \sigma_i^2)$.
  - The two samples are independent.
- Hypothesis:

$$\left.\begin{array}{l} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{array}\right\} \quad \Leftrightarrow \quad \left\{\begin{array}{l} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{array}\right.$$

- Procedure:
  - Calculate

$$\overline{x}_{1,\cdot} = \frac{1}{n_1}\sum_{j=1}^{n_1} x_{1,j} \qquad \text{og} \qquad \overline{x}_{2,\cdot} = \frac{1}{n_2}\sum_{j=1}^{n_2} x_{2,j}$$

  - Is $\overline{x}_{1,\cdot} - \overline{x}_{2,\cdot}$ close to 0 in appropriate distribution?
  - Assumptions gives "appropriate distribution":

$$T = \frac{\overline{x}_{1,\cdot} - \overline{x}_{2,\cdot}}{\widehat{\sigma}_{\overline{x}_{1,\cdot}-\overline{x}_{2,\cdot}}} \sim t\text{-distributed.}$$

- $\widehat{\sigma}_{\overline{x}_{1,\cdot}-\overline{x}_{2,\cdot}}$ and degrees of freedom both depend on the variances!

- Procedure:
  - Calculate

$$\overline{x}_{1,\cdot} = \frac{1}{n_1}\sum_{j=1}^{n_1} x_{1,j} \qquad \text{og} \qquad \overline{x}_{2,\cdot} = \frac{1}{n_2}\sum_{j=1}^{n_2} x_{2,j}$$

  - Is $\overline{x}_{1,\cdot} - \overline{x}_{2,\cdot}$ close to 0 in appropriate distribution?
  - Assumptions gives "appropriate distribution":

$$T = \frac{\overline{x}_{1,\cdot} - \overline{x}_{2,\cdot}}{\widehat{\sigma}_{\overline{x}_{1,\cdot}-\overline{x}_{2,\cdot}}} \sim t\text{-distributed.}$$

- $\widehat{\sigma}_{\overline{x}_{1,\cdot}-\overline{x}_{2,\cdot}}$ and degrees of freedom both depend on the variances!

- Procedure:
  - Calculate

$$\overline{x}_{1,\cdot} = \frac{1}{n_1}\sum_{j=1}^{n_1} x_{1,j} \qquad \text{og} \qquad \overline{x}_{2,\cdot} = \frac{1}{n_2}\sum_{j=1}^{n_2} x_{2,j}$$

  - Is $\overline{x}_{1,\cdot} - \overline{x}_{2,\cdot}$ close to 0 in appropriate distribution?
  - Assumptions gives "appropriate distribution":

$$T = \frac{\overline{x}_{1,\cdot} - \overline{x}_{2,\cdot}}{\widehat{\sigma}_{\overline{x}_{1,\cdot} - \overline{x}_{2,\cdot}}} \sim t\text{-distributed.}$$

- $\widehat{\sigma}_{\overline{x}_{1,\cdot} - \overline{x}_{2,\cdot}}$ and degrees of freedom both depend on the variances!

Paired vs independent test

- We must know from experiment, if samples are paired.
  - Same number of observations, doesn't necessarily mean we can use paired test!
  - If possible, we prefer paired test.

χ^2 test for goodness-of-fit

- Data:

					Total
Class	1	2	...	k	
Observation	o_1	o_2	...	o_k	n
Expected observation	e_1	e_2	...	e_k	n

- Hypothesis:

$$H_0 : \text{Data follows certain distribution}$$
$$H_1 : \text{Data doesn't follow this distribution}$$

- Under H_0: o's $\approx$ e's.
- Test statistic:

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \sim \chi^2(k - m - 1)$$

- Large values of χ^2 are critical for H_0.

χ^2 test for goodness-of-fit

- Data:

					Total
Class	1	2	...	k	
Observation	o_1	o_2	...	o_k	n
Expected observation	e_1	e_2	...	e_k	n

- Hypothesis:

$$H_0 : \text{Data follows certain distribution}$$
$$H_1 : \text{Data doesn't follow this distribution}$$

- Under H_0: o's $\approx$ e's.
- Test statistic:

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \sim \chi^2(k - m - 1)$$

- Large values of χ^2 are critical for H_0.

χ^2 test for goodness-of-fit

- Data:

					Total
Class	1	2	...	k	
Observation	o_1	o_2	...	o_k	n
Expected observation	e_1	e_2	...	e_k	n

- Hypothesis:

H_0 : Data follows certain distribution

H_1 : Data doesn't follow this distribution

- Under H_0: o's $\approx$ e's.
- Test statistic:

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \sim \chi^2(k - m - 1)$$

- Large values of χ^2 are critical for H_0.

χ^2 test for goodness-of-fit

- Data:

					Total
Class	1	2	...	k	
Observation	o_1	o_2	...	o_k	n
Expected observation	e_1	e_2	...	e_k	n

- Hypothesis:

$$H_0 : \text{Data follows certain distribution}$$

$$H_1 : \text{Data doesn't follow this distribution}$$

- Under H_0: o's $\approx$ e's.
- Test statistic:

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \sim \chi^2(k - m - 1)$$

- Large values of χ^2 are critical for H_0.

χ^2 test for goodness-of-fit

- Data:

Class	1	2	...	k	Total
Observation	o_1	o_2	...	o_k	n
Expected observation	e_1	e_2	...	e_k	n

- Hypothesis:

H_0 : Data follows certain distribution

H_1 : Data doesn't follow this distribution

- Under H_0: o's $\approx$ e's.
- Test statistic:

num. of parameters, estimated

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \sim \chi^2(k - m - 1)$$

- Large values of χ^2 are critical for H_0.

χ^2 test for goodness-of-fit

- Data:

					Total
Class	1	2	...	k	
Observation	o_1	o_2	...	o_k	n
Expected observation	e_1	e_2	...	e_k	n

- Hypothesis:

H_0 : Data follows certain distribution

H_1 : Data doesn't follow this distribution

- Under H_0: o's $\approx$ e's.
- Test statistic:

num. of parameters, estimated

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \sim \chi^2(k - m - 1)$$

- Large values of χ^2 are critical for H_0.

χ^2 test for goodness-of-fit

- Data:

					Total
Class	1	2	...	k	
Observation	o_1	o_2	...	o_k	n
Expected observation	e_1	e_2	...	e_k	n

- Hypothesis:

H_0 : Data follows certain distribution

H_1 : Data doesn't follow this distribution

- Under H_0: o's $\approx$ e's.
- Test statistic:

num. of parameters, estimated

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \sim \chi^2(k - m - 1)$$

- Large values of χ^2 are critical for H_0.

χ^2 test example

- A die is rolled 600 times:

1	2	3	4	5	6
119	117	98	89	96	81

- Hypothesis:

$$H_0 : \text{The die is fair}$$
$$H_1 : \text{The die is not fair}$$

- Under H_0 the expected table is:

1	2	3	4	5	6
100	100	100	100	100	100

- Calculate test statistic:

$$\chi^2 = \frac{(119-100)^2}{100} + \frac{(117-100)^2}{100} + \cdots + \frac{(81-100)^2}{100} \approx 11.5$$

- Calculate p-value in $\chi^2(5)$ distribution $= 4.2\%$

χ^2 test example

- A die is rolled 600 times:

1	2	3	4	5	6
119	117	98	89	96	81

- Hypothesis:

$$H_0 : \text{The die is fair}$$
$$H_1 : \text{The die is not fair}$$

- Under H_0 the expected table is:

1	2	3	4	5	6
100	100	100	100	100	100

- Calculate test statistic:

$$\chi^2 = \frac{(119-100)^2}{100} + \frac{(117-100)^2}{100} + \cdots + \frac{(81-100)^2}{100} \approx 11.5$$

- Calculate p-value in $\chi^2(5)$ distribution $= 4.2\%$

χ^2 test example

- A die is rolled 600 times:

1	2	3	4	5	6
119	117	98	89	96	81

- Hypothesis:

$$H_0 : \text{The die is fair}$$
$$H_1 : \text{The die is not fair}$$

- Under H_0 the expected table is:

1	2	3	4	5	6
100	100	100	100	100	100

- Calculate test statistic:

$$\chi^2 = \frac{(119-100)^2}{100} + \frac{(117-100)^2}{100} + \cdots + \frac{(81-100)^2}{100} \approx 11.5$$

- Calculate p-value in $\chi^2(5)$ distribution $= 4.2\%$

χ^2 test example

- A die is rolled 600 times:

1	2	3	4	5	6
119	117	98	89	96	81

- Hypothesis:

$$H_0 : \text{The die is fair}$$
$$H_1 : \text{The die is not fair}$$

- Under H_0 the expected table is:

1	2	3	4	5	6
100	100	100	100	100	100

- Calculate test statistic:

$$\chi^2 = \frac{(119-100)^2}{100} + \frac{(117-100)^2}{100} + \cdots + \frac{(81-100)^2}{100} \approx 11.5$$

- Calculate p-value in $\chi^2(5)$ distribution $= 4.2\%$

Remarks about χ^2 tests

- NOTE: The χ^2 is an approximate distribution for the test statistic and the approximation is typically not good if any expected count is less than 5.
- "χ^2 test" cover several tests:
  - Test distribution.
  - Test for independence between observations in table.
  - Test for homogenity in table.
- Common features:
  - Data is in table.
  - Test statistic:

$$\sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}.$$

  - Test statistic is χ^2 distributed.

Remarks about χ^2 tests

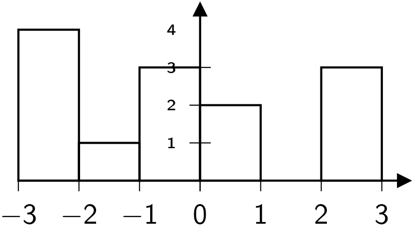
- NOTE: The χ^2 is an approximate distribution for the test statistic and the approximation is typically not good if any expected count is less than 5.
- "χ^2 test" cover several tests:
  - Test distribution.
  - Test for independence between observations in table.
  - Test for homogenity in table.
- Common features:
  - Data is in table.
  - Test statistic:

$$\sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}.$$

  - Test statistic is χ^2 distributed.

Histogram

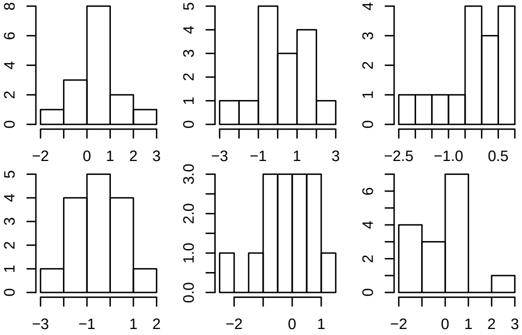
- We observe n values $x_1, x_2, \dots, x_n$.
- Histogram of data:
  - Divide the x axis into equally sized intervals.
  - Make a bar in each interval that represents the number of observations in the interval.



Observations:
-2.9, -2.9, -2.5, -2.1, -1.5, -0.75, -0.5, -0.25, 0.5, 0.5, 2, 2.5, 2.6

Problems with histograms

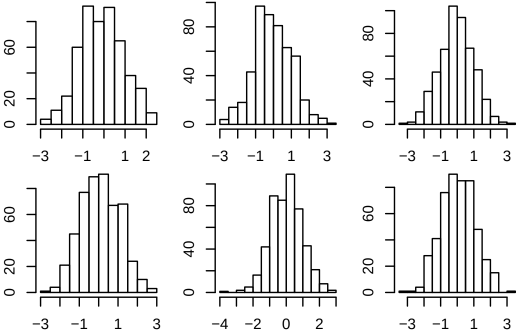
- With few observations it can be hard to recognize the distribution
- The intervals should be chosen apropriately (rarely a serious problem)



20 normal distributed observations in each histogram.

Problems with histograms

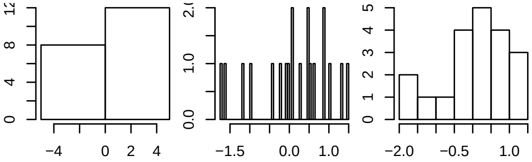
- With few observations it can be hard to recognize the distribution
- The intervals should be chosen apropriately (rarely a serious problem)



200 normal distributed observations in each histogram.

Problems with histograms

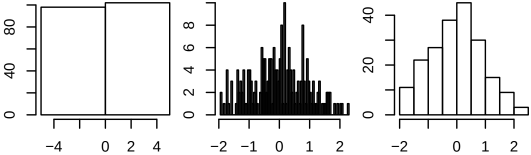
- With few observations it can be hard to recognize the distribution
- The intervals should be chosen apropriately (rarely a serious problem)



20 normal distributed observations in each histogram.

Problems with histograms

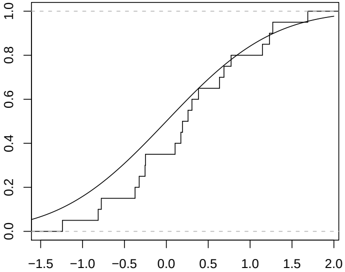
- With few observations it can be hard to recognize the distribution
- The intervals should be chosen apropriately (rarely a serious problem)



200 normal distributed observations in each histogram.

QQ-plot

- QQ-plot: 2. method to explore distribution of data.
- Idea: Compare empirical distribution function with a theoretical distribution function.

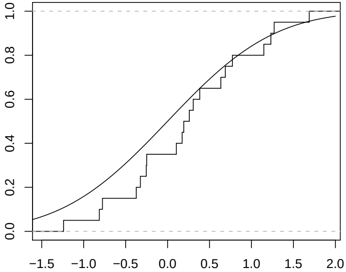


20 normal distributed observations.

- Problem: It is difficult to compare graphs, that are not straight lines.
- New idea: Compare empirical quantiles with theoretical quantiles.

QQ-plot

- QQ-plot: 2. method to explore distribution of data.
- Idea: Compare empirical distribution function with a theoretical distribution function.

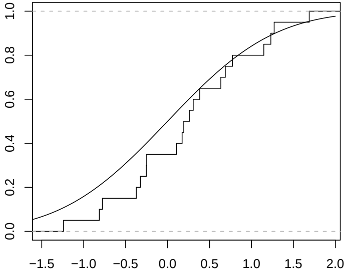


20 normal distributed observations.

- Problem: It is difficult to compare graphs, that are not straight lines.
- New idea: Compare empirical quantiles with theoretical quantiles.

QQ-plot

- QQ-plot: 2. method to explore distribution of data.
- Idea: Compare empirical distribution function with a theoretical distribution function.

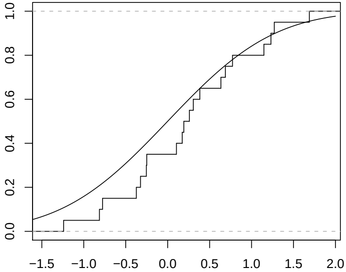


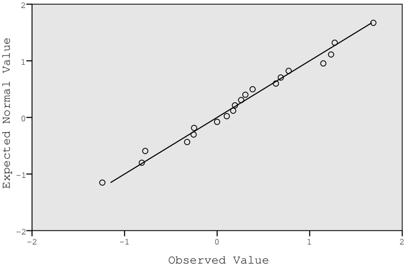
20 normal distributed observations.

- Problem: It is difficult to compare graphs, that are not straight lines.
- New idea: Compare empirical quantiles with theoretical quantiles.

QQ-plot

- QQ-plot: 2. method to explore distribution of data.
- Idea: Compare empirical distribution function with a theoretical distribution function.



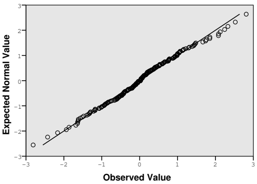


20 normal distributed observations.

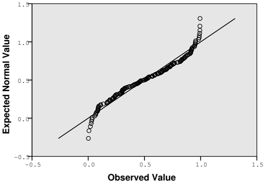
- Problem: It is difficult to compare graphs, that are not straight lines.
- New idea: Compare empirical quantiles with theoretical quantiles.

QQ-plots

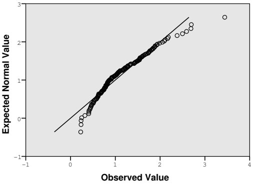
- Facts about QQ-plots
  - If the proposed distribution is good, we get a linear pattern.
  - As the number of observations increase, the variability decreases.
  - The points are dependent, so they have a tendency to twist around the line.
- Examples of QQ-plots



Normal distributed data



Uniformly distributed data



Lognormal distributed data