Hypothesis tests
Hypothesis

H_0: Null-hypothesis is an conjecture which we assume is true until we have too much evidence against it.

H_1: The alternative hypothesis covers the alternative to H_0

Notice: We either reject or cannot reject the null-hypothesis – we *never* accept (or prove) the null-hypothesis.

Legal analogy:

H_0: Innocent vs H_1: Guilty

Procedure: Defendant is assumed innocent (H_0) until enough evidence suggest he is guilty..

Hypothesis tests
Examples of hypotheses

The mean height of men is 1.80 cm:

$$H_0 : \mu = 180$$
$$H_1 : \mu \neq 180$$

The mean age of when people leave home is higher than 18 years:

$$H_0 : \mu \geq 18$$
$$H_1 : \mu < 18$$

The average price of milk in 7-11 is 2 kr higher than in Brugsen:

$$H_0 : \mu_{7-11} - \mu_b = 2$$
$$H_1 : \mu_{7-11} - \mu_b \neq 2$$

Notice: H_0 always involves equality (≤,≥, or =)

Hypothesis tests
One-sided and two-sided tests

One-sided test:

$$H_0 : \theta \geq \theta_0 \qquad H_0 : \theta \leq \theta_0$$

$$H_1 : \theta < \theta_0 \qquad H_1 : \theta > \theta_0$$

Two-sided test:

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta \neq \theta_0$$

Hypothesis tests
Test statistics

Recall: A sample function is a function of data

Test statistic: a sample function T, which indicates if the null hypothesis should be rejected or not.

Critical area: if T lies in the critical area then the null-hypothesis is rejected.

Critical values: boundary points for the critical area.

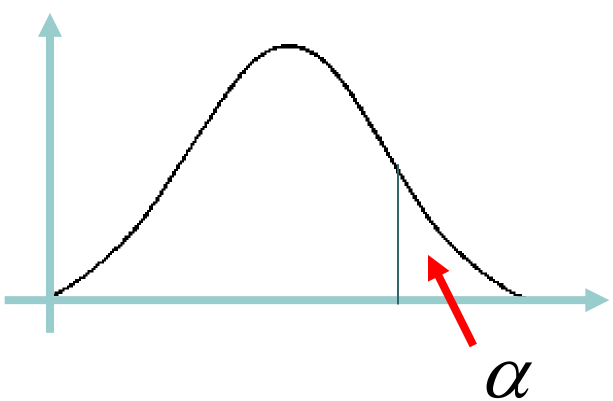
Hypothesis tests
Type I and type II errors

Type I error: H_0 rejected, when H_0 is true.

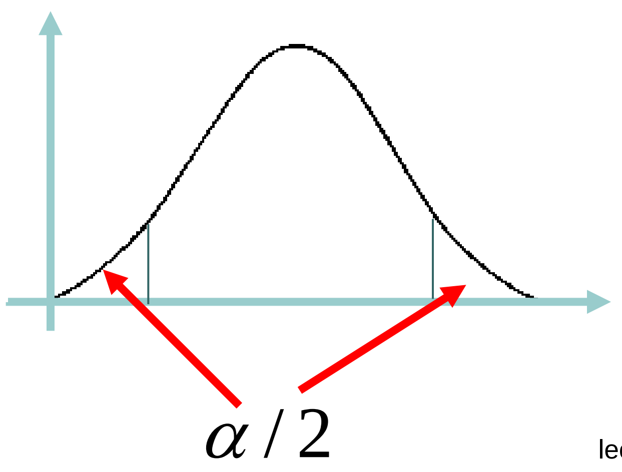
Type II error: H_0 not rejected, when H_0 is false.

Significance level: α is the probability of committing a Type I error.

One-sided test



Two-sided test



Test of mean
Variance known (two-sided)

Hypothesis:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

Significance level:

$$P\left(-z_{\alpha/2} < \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} < z_{\alpha/2} \right) = 1 - \alpha$$

Test statistics:

$$z = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$$

Critical values:

$$-z_{\alpha/2}, z_{\alpha/2}$$

Decision: reject H_0 if z does not lie between the critical values, otherwise we cannot reject H_0.

Test of mean
Variance known

Problem: Anders claims that a flight to Paris on average costs 5000 kr. For a sample of ten flights to Paris he finds a sample average of 4850kr.

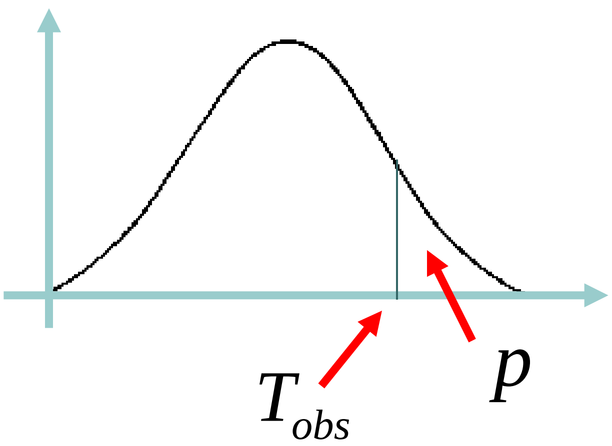
1. Rephrase Anders' claim as a statistical hypothesis.
2. Assume that the standard deviation of prices of flights to Paris is 100 kr. Is Anders' hypothesis reasonable at the 5% significance level?

Hypothesis tests
p-value

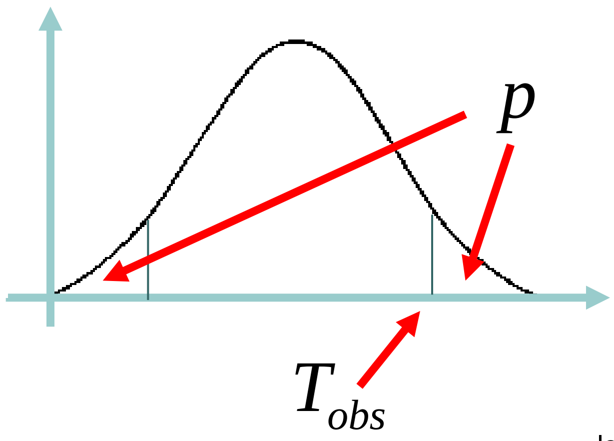
Assuming that H_0 is true, the p-value is the probability of observing a more extreme test statistics than the one just observed.

The null-hypothesis if rejected if *p*-value < α.

One-sided test



Two-sided test



Test of mean
Variance known

Problem (cont.): Anders claims that a flight to Paris on average costs 5000 kr. For a sample of ten flights to Paris he finds a sample average of 4850kr.

3. Calculate the *p*-value and compare it to the significance level (we still assume the standard deviation to be 100kr).

Hypothesis test and confidence intervals

The connection with confidence intervals:

$$-z_{\alpha/2} < \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} < z_{\alpha/2}$$

$$\Updownarrow$$

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu_0 < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

That is, the null-hypothesis, $\mu = \mu_0$, is rejected if and only if μ_0 lies outside the confidence interval.

Test of mean
Variance known (one-sided)

Hypotheses:

$$H_0 : \mu \geq \mu_0$$
$$H_1 : \mu < \mu_0$$

Significance level:

$$P\left(-z_\alpha < \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \right) = 1 - \alpha$$

Test statistic:

$$z = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$$

Critical value:

$$-z_\alpha$$

Decision: Reject H_0 if z lies below the critical value.

Test of mean
Variance unknown (two-sided)

Hypotheses:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

Significance level:

$$P\left(-t_{\alpha/2,n-1} < \frac{\overline{X} - \mu_0}{s/\sqrt{n}} < t_{\alpha/2,n-1} \right) = 1 - \alpha$$

Test statistic:

$$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}}$$

Critical values:

$$-t_{\alpha/2,n-1}, t_{\alpha/2,n-1}$$

Decision: Reject H_0 if z does not lie between the critical values.

Test of mean the MATLAB way Variance unknown (two-sided)

Default: H_0: $\mu = \mu_0$

Data

$\mu_0 = 5000$

```
>> [h,p,ci,stats]=ttest(x,5000,0.01)

h =
     1
p =
    0.0013
ci =
  1.0e+003 *
    4.7436
    4.9564
stats =
    tstat: -4.5804
       df: 9
       sd: 103.5584
```

Significance level $\alpha = 0.01$

h = 0 : H_0 not rejected
h = 1 : H_0 rejected

p-value

$(1-\alpha)100\%$ confidence interval

tstat = $\dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

df = degrees of freedom

sd = sample standard deviation

Test of mean the R way
Variance unknown (two-sided)

Data

$\mu_0 = 5000$

```
> t.test(x=x,mu=5000,conf.level=0.99)

        One Sample t-test

data:  x
t = -4.5804, df = 9, p-value = 0.001327
alternative hypothesis: true mean is not equal to 5000
99 percent confidence interval:
 4743.574 4956.426
sample estimates:
mean of x
     4850
```

Confidence level: $1-\alpha = 0.99$

Test of two means
Known variances (two-sided)

Hypotheses:

$$H_0 : \mu_1 - \mu_2 = d_0$$
$$H_1 : \mu_1 - \mu_2 = d_0$$

Significance level:

$$P\left(-z_{\alpha/2} < z < z_{\alpha/2}\right) = 1 - \alpha$$

Test statistic:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}}$$

Critical values:

$$-z_{\alpha/2}, z_{\alpha/2}$$

Decision: Reject H_0 if z does not lie between the critical values

Test of two means

Unknown & equal variances (two-sided)

Test statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p \sqrt{1/n_1 + 1/n_2}}$$

Pooled variance estimate:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Critical values:

$$-t_{\alpha/2,v}, t_{\alpha/2,v}$$

Degrees of freedom

$$v = n_1 + n_2 - 2$$

Test of two means

Unknown & unequal variances (two-sided)

Test statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{s_1^2 / n_1 + s_2^2 / n_2}}$$

Critical values:

$$-t_{\alpha/2,v}, t_{\alpha/2,v}$$

Degrees of freedom:

$$v = \frac{(s_1^2 / n_1 + s_2^2 / n_2)^2}{\frac{(s_1^2 / n_1)^2}{n_1 - 1} + \frac{(s_2^2 / n_2)^2}{n_2 - 1}}$$

Test of two means the MATLAB way Equal unknown variance (two-sided)

```
>> [h,p,ci,stats] = ttest2(x,y,0.05,'both','equal')

h =
     0
p =
    0.0715
ci =
   -0.0490
    1.0490
stats =
    tstat: 1.9305
       df: 16
       sd: 0.5460
```

Data (x, y)

H_0: $\mu_1 - \mu_2 = 0$ ('both')

$\sigma^2_1 = \sigma^2_2$ ('equal')

Significance level $\alpha = 0.05$

h = 0 : H_0 not rejected

h = 1 : H_0 rejected

p-value

$(1-\alpha)100\%$ confidence interval for $\mu_1 - \mu_2$

$$\text{tstat} = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p\sqrt{1/n_1 + 1/n_2}}$$

df = degrees of freedom

$$\text{sd}^\wedge 2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$$

Hypothesis tests
A couple of remarks

Many tests exist in both a two-sided and a one-sided versions:

- Null-hypotheses: ≤, =, or ≥
- Critical values: one-sided use α
  two-sided use $\alpha/2$

We can reject H_0 in three equivalent way:

- The test statistic is in the critical area
- The *p*-value < α
- Hypothesised value (e.g. μ_0) is outside the confidence intervals (only for two-sided tests)

Test of variance
Two-sided test

Hypotheses:

$$H_0 : \sigma = \sigma_0$$
$$H_1 : \sigma \neq \sigma_0$$

Significance level:

$$P\left(\chi^2_{1-\alpha/2} < \frac{(n-1)s^2}{\sigma_0^2} < \chi^2_{\alpha/2} \right) = 1 - \alpha$$

Test statistic:

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Critical values:

$$\chi^2_{1-\alpha/2}, \chi^2_{\alpha/2}$$

Decision: Reject H_0 if χ^2 does not lie between the critical values

Test of variance

Problem: A company produces golf balls, where the standard deviation of the diameter is not allowed to exceed 0.5 mm. Bente wants to prove that the variation is higher, and so measures 10 golf balls. The sample standard deviation is 0.6 mm.

Is Bente's claim correct?
Formulate the hypotheses and test it the using a 5% significance level.

Test of variance
The MATLAB way

Data

$\sigma^2_0 = 0.5^2$

H_0: $\sigma^2 \le \sigma^2_0$

```
>> [h,p,ci,stats]=vartest(x,0.5^2,0.05,'right')

h =
     0
p =
    0.1644
ci =
    0.1915
       Inf
stats =
    chisqstat: 12.9600
           df: 9
```

Significance level $\alpha = 0.05$

h = 0 : H_0 not rejected
h = 1 : H_0 rejected

p-value

$(1-\alpha)100\%$ confidence interval

chisqstat $= \frac{(n-1)s^2}{\sigma^2}$

df = degrees of freedom

Ratio of variances
Estimation

Consider two populations with variances σ^2_1 and σ^2_2 and two samples of sizes n_1 and n_2.

Estimate for σ^2_1 / σ^2_2 :

$$s_1^2 / s_2^2$$

Confidence interval:

$$\frac{s_1^2}{s_2^2} \frac{1}{f_{\alpha/2, n_1-1, n_2-1}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} f_{\alpha/2, n_2-1, n_1-1}$$

Test of two variances
Two-sided test

Hypotheses:

$$H_0 : \sigma_1 = \sigma_2$$
$$H_1 : \sigma_1 \neq \sigma_2$$

Significance level:

$$P\left(1/\, f_{\alpha/2, n_2-1, n_1-1} < f < f_{\alpha/2, n_2-1, n_2-1}\right) = 1-\alpha$$

Test statistic:

$$f = \frac{s_1^2}{s_2^2}$$

Critical values:

$$1/\, f_{\alpha/2, n_2-1, n_1-1},\ f_{\alpha/2, n_1-1, n_2-1}$$

Decision: Reject H_0 if f does not lie between the critical values.

Test of equal variance

Problem: Bente now wants to compare two companies to see if one company makes more equally sized golf ball that the other. For company A she obtains a sample std. div. of 0.5 mm for 10 golf balls, and for company B she obtains a sample std. div. of 0.6 mm for 8 golf balls.

Is there a significant difference? Test this using a 10% significance level.

Test of equal variance
The MATLAB way

Data pop. A

Data pop. B

```
>> [h,p,ci,stats] = vartest2(x,y,0.1)

h =
     0
p =
    0.5972
ci =
    0.1889
    2.2866
stats =
    fstat: 0.6944
      df1: 9
      df2: 7
```

Significance level $\alpha = 0.1$

h = 0 : H_0 not rejected
h = 1 : H_0 rejected

p-value

$(1-\alpha)100\%$ confidence interval

$$\text{fstat} = \frac{s_1^2}{s_2^2}$$

df1 and df2 = degrees of freedom

A couple of comments

- Overview of tests of means on page 351
- In the next and final lecture we'll have a brief look a Bayesian statistics – a different way of doing statistics.
- Next time the exercises will consist in you having a go at some old exam problems.