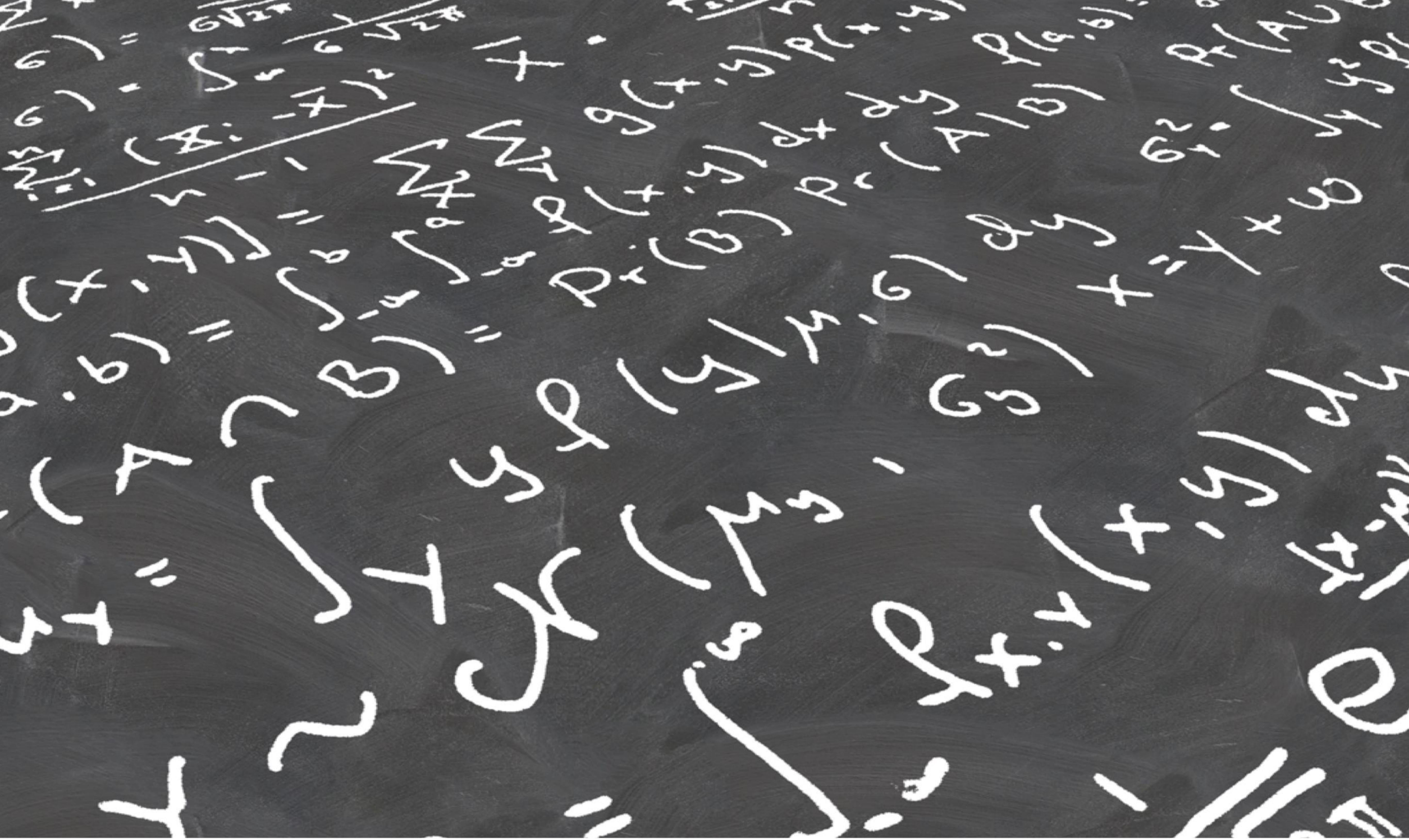


10. Hypothesis Test

Gunvor Elisabeth Kirkelund
Lars Mandrup
Slides and material provided in parts by
Henrik Pedersen

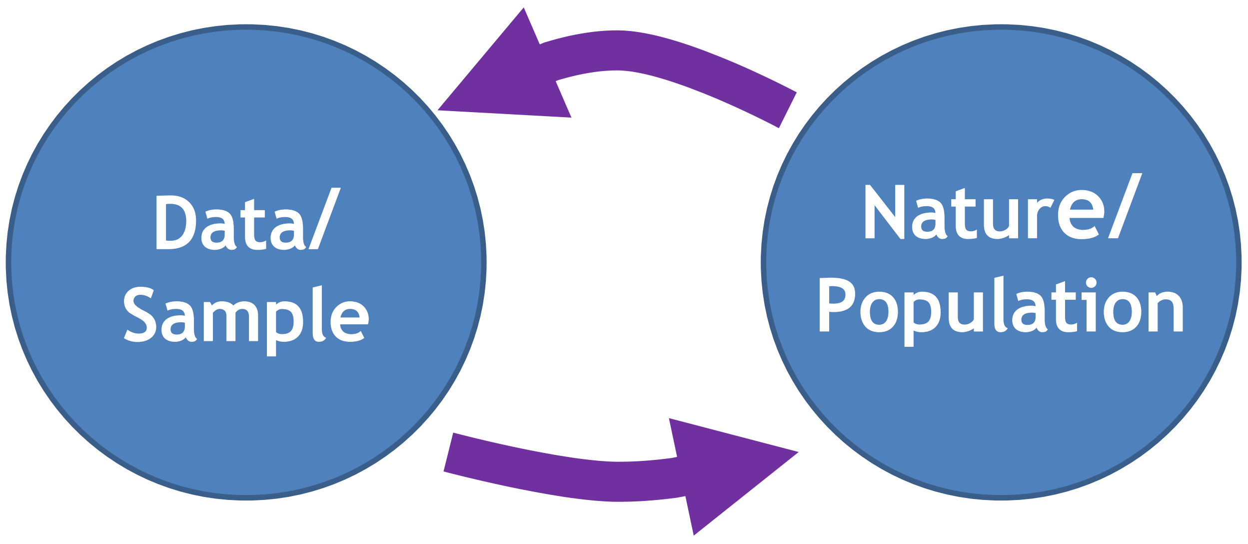
Todays Content

- Repetition from last time
- Hypothesis Test
- Test of the mean with known variance (z-test)
- Test of the mean with unknown variance (t-test)
- Q-Q test for normality

Introduction to Statistics

Probability theory

Given the cause (population), what should the data (sample) look like?



Statistics

Given the data (sample), what caused them (population)?

- Testing a hypotesis
- Estimating means and variances
- If we don’t know better: We assume data are normally distributed

Statistical Model

Statistical model:

- A random sample and its pdf, $f_X(x;\theta)$, where θ is the parameter(s) of the pdf.

- Because of the Central Limit Theorem (CLT) we often can use the normal distribution $\mathcal{N}(\mu, \frac{\sigma^2}{n})$ with mean μ and variance σ^2/n as statistical model for the sample mean $\bar{X}$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(\mu, \frac{\sigma^2}{n}) \quad \text{(n>30)}$$

Estimator

Estimator:

- An estimator $\hat{\theta}(X)$ is a statistic used to estimate the unknown parameter θ of a random sample X.
- An estimator is unbiased if $E\left[\hat{\theta}\right] = \theta$.

Unbiased estimators:

- The sample mean: $\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Degrees of freedom

- The sample variance: $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

Degrees of freedom

- The sample succes rate: $\hat{p} = \frac{x}{n}$

Number of successes

Number of trials

- The sample event rate: $\hat{\lambda} = \frac{x}{t}$

Number of events

Time

Test Statistics

Test statistics:

- A random variable that summarized a data-set by reducing the data to one value that can be used to perform the hypothesis test.

Fx n>30

- For a sample assumed to follow the normal distribution $\mathcal{N}(\mu, \sigma^2)$ with **known mean** $\boldsymbol{\mu}$ and **variance** $\boldsymbol{\sigma^2}$ we can use the z-statistics (z-score):

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim \mathcal{N}(0,1)$$

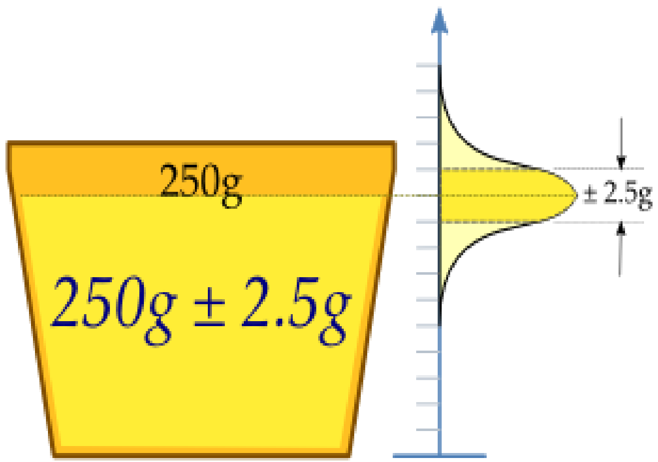
Standard (normalized) normal distribution (μ=0 and σ^2=1)

➢ **Assignment:** Show that if $X \sim \mathcal{N}(\mu, \sigma^2)$ then $Z = \frac{\bar{X}-\mu}{\sigma / \sqrt{n}} \sim \mathcal{N}(0,1)$ (ie. having μ=0 and σ^2=1)

Cup Example

- A machine fills cups with a liquid, the content of the cups is 250 g. of liquid.
- The machine cannot fill with exatly 250 g. The content added to individual cups shows some variation, and is considered a random variable X.

If the machine is adequately calibrated, X is normally distributed with mean μ = 250g and standard variation σ = 2.5g



- Statistical model: $X \sim \mathcal{N}(250, 2.5)$
- Test sample: $X_1, X_2, \dots, X_{25}$
- Statistics: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- Estimator: $\hat{\mu} = \bar{X} = \frac{1}{25}\sum_{i=1}^{25} X_i = 250.2\ g$
- z-score: $z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = 0.4$

(μ = 250g)

➢ Can the statement, that the machine is adequately calibrated be rejected?

Hypothesis

- **Definition – Null hypothesis (H_0)**
  - The statement being tested in a test of statistical significance is called the **null hypothesis**. The test of significance is designed to assess the strength of the evidence against the null hypothesis.
  - Usually, the null hypothesis is a statement of 'no effect‘, 'no difference‘ or ‘no relation' between the phenomena whose relation is under investigation.
- **Definition – Alternative hypothesis (H_1)**
  - The statement that is hoped or expected to be true instead of the null hypothesis is the **alternative hypothesis**
  - The alternative hypothesis, as the name suggests, is the alternative to the null hypothesis: it states that there is some 'effect/difference’ or some ‘kind of relation'.

Important!

- One cannot “prove” a null hypothesis, one can only test how close it is to being true.
- Therefore, we never say that we *accept* the null hypothesis, but that we either ***reject it*** or ***fail to reject it***.

Hypothesis

An example of a null hypothesis:

- A certain drug may reduce the chance of having a heart attack.
- Possible null hypothesis H_0: “This drug has no effect on the chances of having a heart attack".
- An alternative hypothesis H_1: “This drug has an effect on the chances of having a heart attack”.
- The test of the hypothesis consists of giving the drug to half of the people in a study group as a controlled experiment.
- If the data show a statistically significant change in the people receiving the drug, the null hypothesis is rejected.

Hypothesis

- The term "**null hypothesis**" H_0 is a general statement or default position that there is no relationship between two measured phenomena, or no association among groups.
- Rejecting or disproving the null hypothesis is a central task in the modern practice of science; the field of statistics gives precise criteria for rejecting a null hypothesis.
- The null hypothesis H_0 is generally assumed to be true until evidence indicates otherwise.
- A null hypothesis is rejected if the observed data are significantly unlikely to have occurred if the null hypothesis were true. In this case an alternative hypothesis H_1 is accepted in its place – concluding that there are grounds for believing that there *is* a relationship between two phenomena.
- If the data are consistent with the null hypothesis, then the null hypothesis is not rejected (i.e., accepted).
- **In neither case is the null hypothesis or its alternative proven**; the null hypothesis is tested with data and a decision is made based on how **likely or unlikely** the data are. This is analogous to a criminal trial, in which the defendant is assumed to be innocent (null is not rejected) until proven guilty (null is rejected) beyond a reasonable doubt (to a statistically significant degree).

Hypothesis testing

- Hypothesis testing works by collecting a randomly selected representative sample X (data) and measuring how likely the particular set of data is, assuming the null hypothesis H_0 is true: $Pr(X|H_0)$

- The data-set is usually specified via a **test statistic** which is designed to measure the extent of apparent departure from the null hypothesis – fx. z-statistic or t-statistic.

- If the data-set of a randomly selected representative sample is very unlikely relative to the null hypothesis – i.e. only rarely (usually in less than either 5% or 1% (the significance level α)) will be observed – we rejects the null hypothesis concluding it (probably) is false.

- If the data do not contradict the null hypothesis, then only a weak conclusion can be made: namely, that the observed data set provides no strong evidence against the null hypothesis. In this case, because the null hypothesis could be true or false, it is interpreted as there is no evidence to support changing from a currently useful regime (the null hypothesis) to a different one.

Cup Example

- **Example 1 – cup filling example**
  - If the machine is adequately calibrated, the true population mean should be 250 grams. Hence, the null hypothesis is

$$H_0\colon\ \mu = 250$$

  - If we are not concerned about the direction of a possible deviation from $\mu = 250$, the alternative hypothesis is

$$H_1\colon\ \mu \neq 250$$

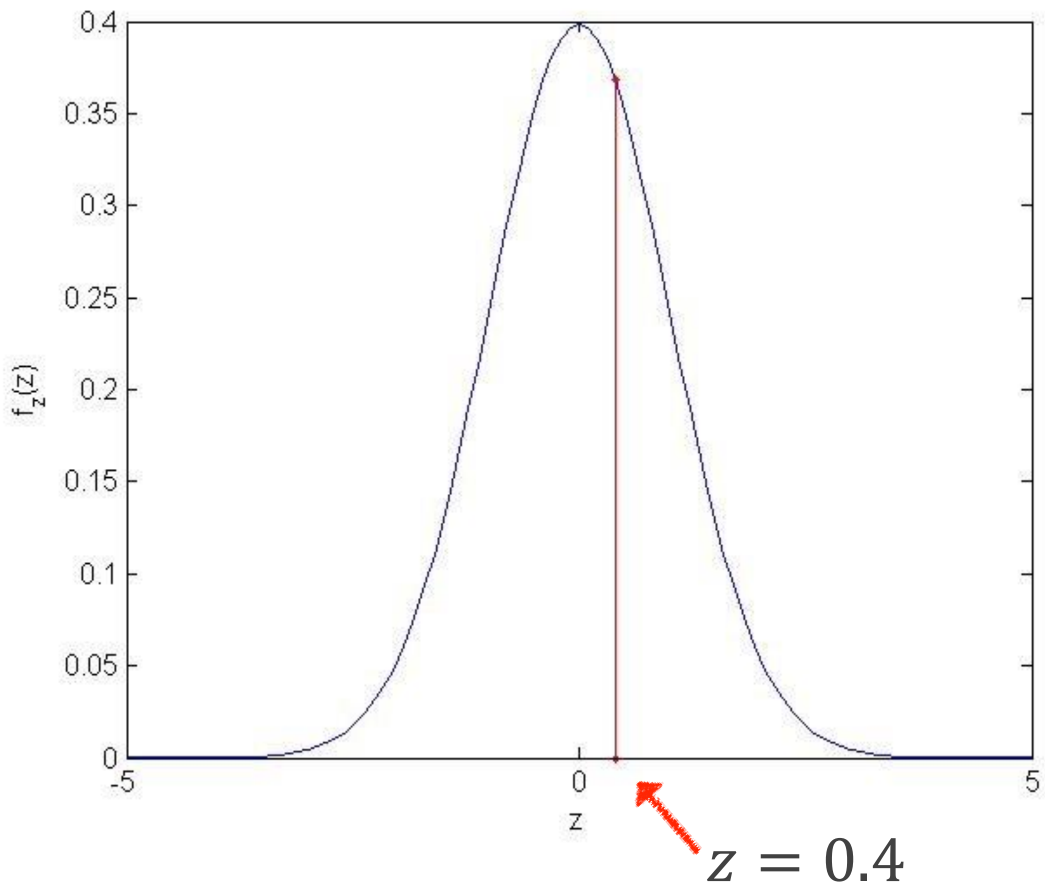
Cup Example

Since we know both the true mean μ (the hypothesis) and variance σ^2, we use the test statistics z

Test statistics: $z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$

z-score: $z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{250.2-250}{2.5/\sqrt{25}} = 0.4$

Standard normal distribution (PDF)



Does it seem plausible that z=0.4 is an observation drawn from a standard normal distribution?

Same as asking: what is the probability of observing a test size (z) that is more extreme than 0.4?

p-value

The lowest significance level
α that results in rejecting
the null hypothesis

p-value

- The p-value is the probability of getting a result equal to or more extreme than the observed test-sample X under the assumption of a null hypothesis H_0:

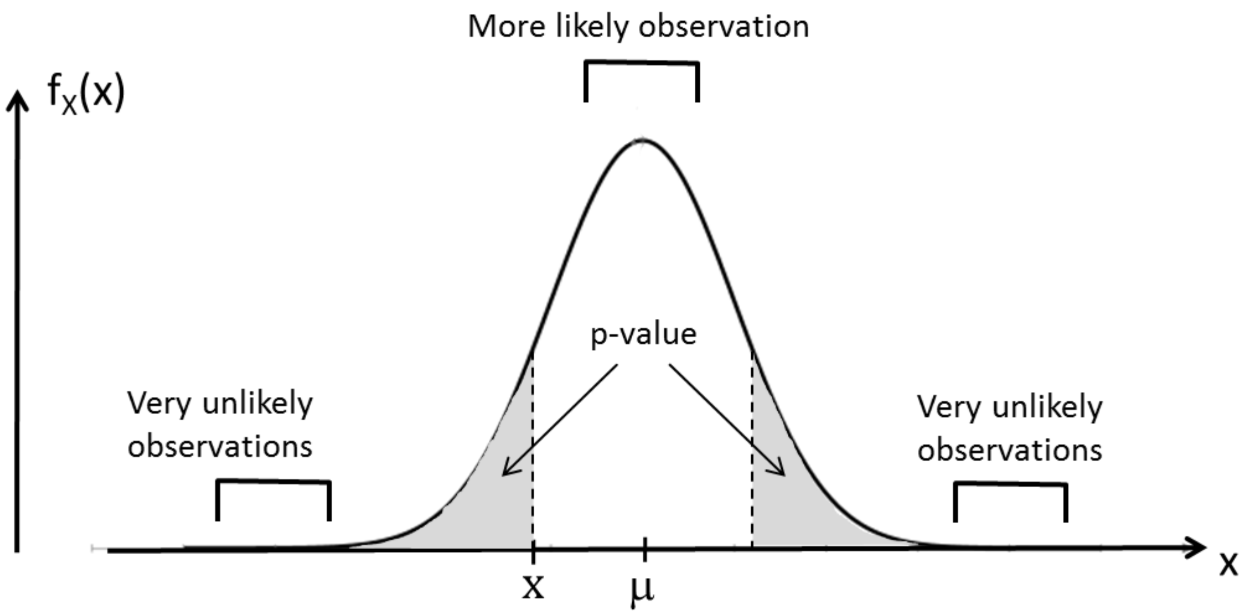
$$p\text{-}value = Pr(Worse\ result\ than\ X | H_0)$$

- If $\boldsymbol{x}$ denotes the observed quantity, the p-value is:
  - $\boldsymbol{Pr(X \geq x | H_0)}$ for a right-tailed event
  - $\boldsymbol{Pr(X \leq x | H_0)}$ for a left-tailed event
  - $2 \cdot min\{\boldsymbol{Pr(X \leq x | H_0)}, \boldsymbol{Pr(X \geq x | H_0)}\}$ for a two-tailed event

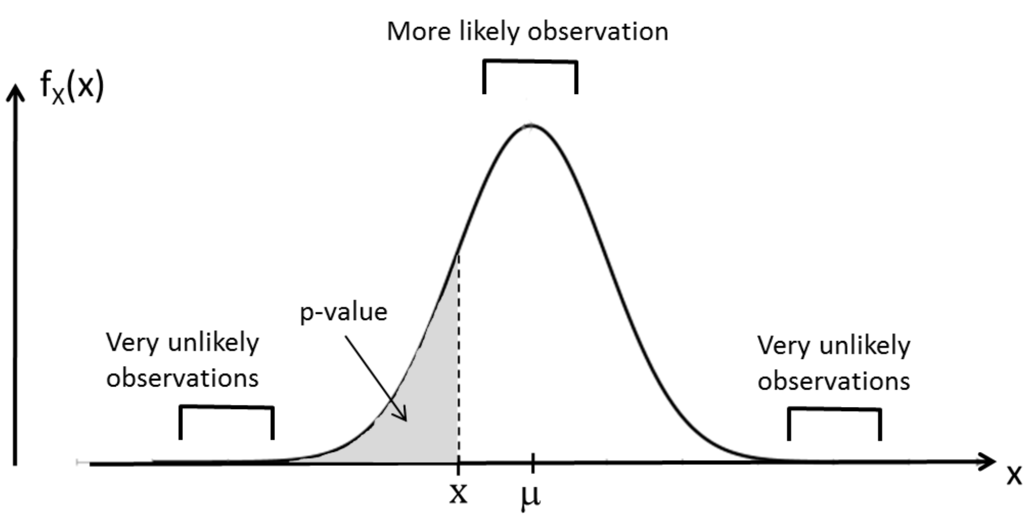
p-value

Two-tailed $p = 2 \cdot min\{\boldsymbol{Pr(X \leq x|H_0)}, \boldsymbol{Pr(X \geq x|H_0)}\}$

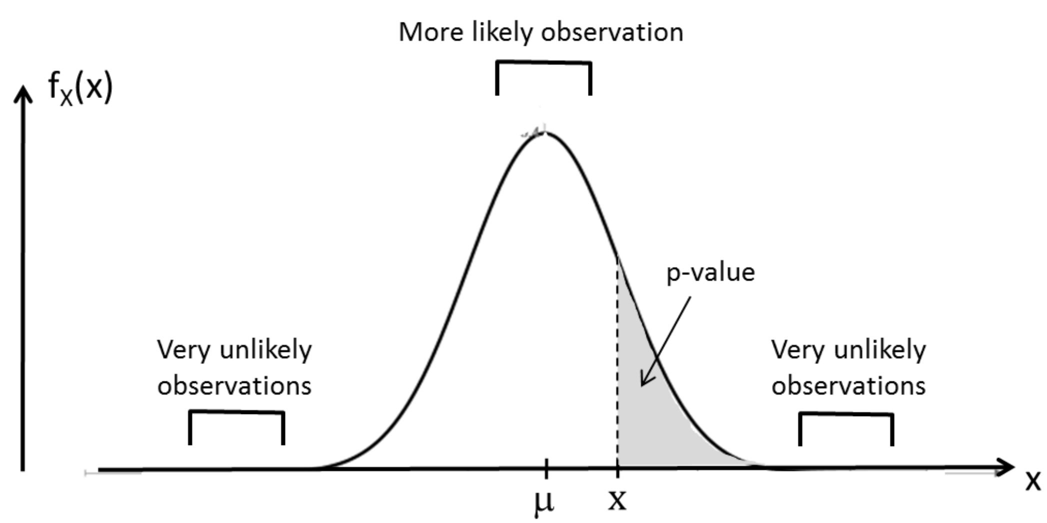
$$= Pr(X - \mu \leq -|x - \mu|) + Pr(X - \mu \geq |x - \mu|)$$



Left-tailed $\boldsymbol{p = Pr(X \leq x|H_0)}$



Right-tailed $\boldsymbol{p = Pr(X \geq x|H_0)}$



Significance Level

Significance level α

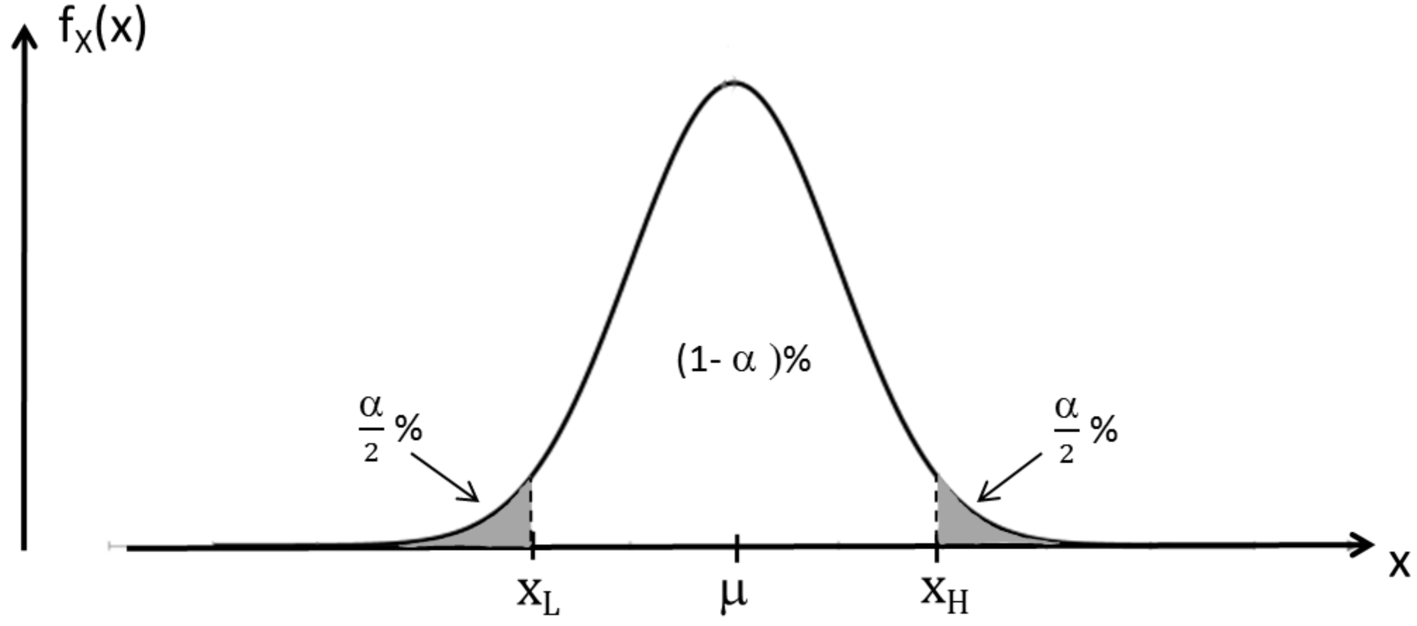
- The statistical significance level α is the lower limit we will accept for the probability of getting a more extreme result assuming the null hypothesis H_0 is true (p-value).

- By comparing the p-value for the test with the significance level α we can deside whether the null hypothesis H_0 should be rejected (p-value $< \alpha$) or we fail to reject it (p-value $> \alpha$).

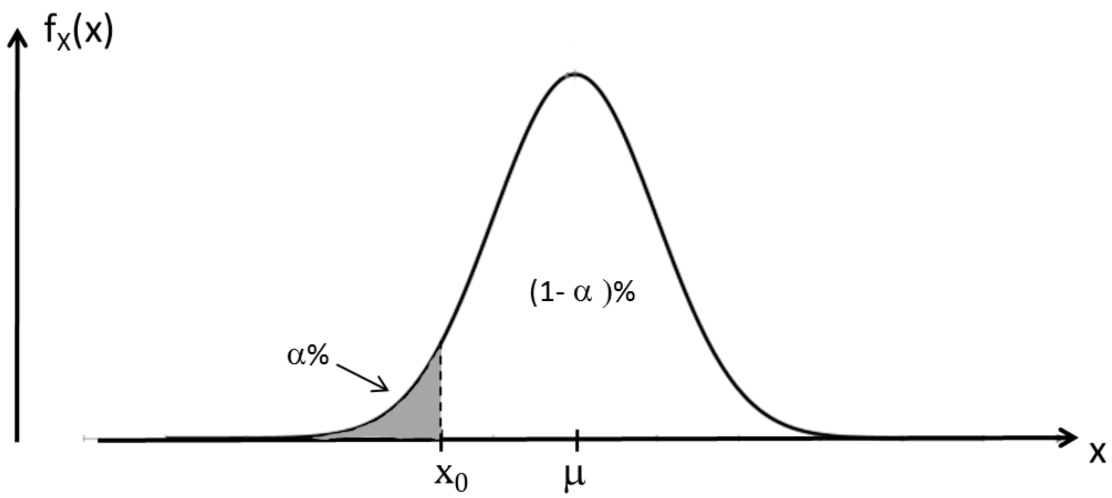
- The most common used significance level is $\alpha = 0{,}05$ (5%)

Significance Level

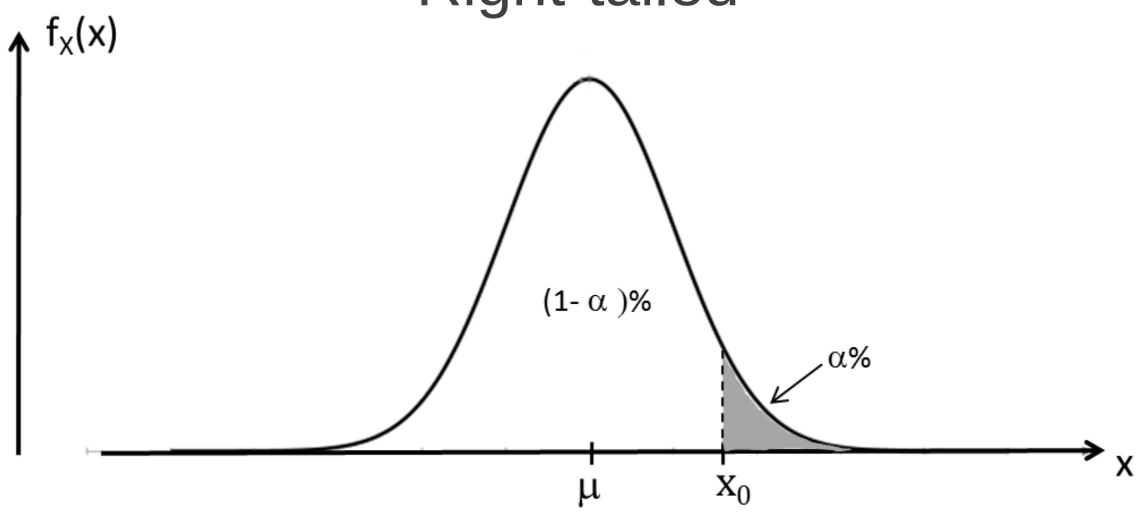
Two-tailed



Left-tailed



Right-tailed



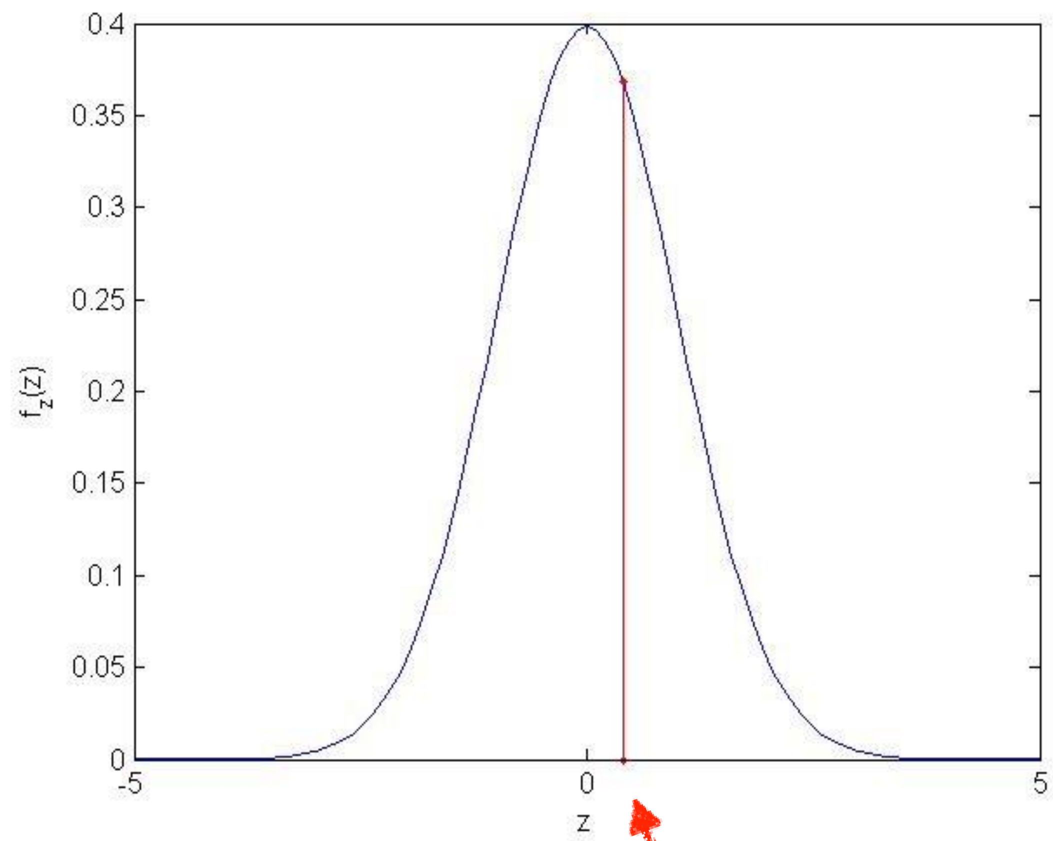
Cup Example

p-value:

$$Pr(Z > z \cup Z < -z)$$
$$= Pr(Z > z) + Pr(Z < -z)$$
$$= 2 \cdot Pr(Z > z)$$
$$= 2 \cdot (1 - \Pr(Z \leq z))$$
$$= 2 \cdot (1 - \Pr(Z \leq 0.4))$$
$$= 2 \cdot \left(1 - \Phi(0.4)\right) \leftarrow \texttt{normcdf(0.4)}$$
$$= 2 \cdot (1 - 0.6554)$$
$$= 0.6892 \rightarrow \text{Compare with } \alpha = 0.05$$

Here, we fail to reject the null hypothesis (H_0: $\mu = 250$), because the p-value is larger than $\alpha = 0.05$.

Standard normal distribution (PDF)



z-score: $z = 0.4$

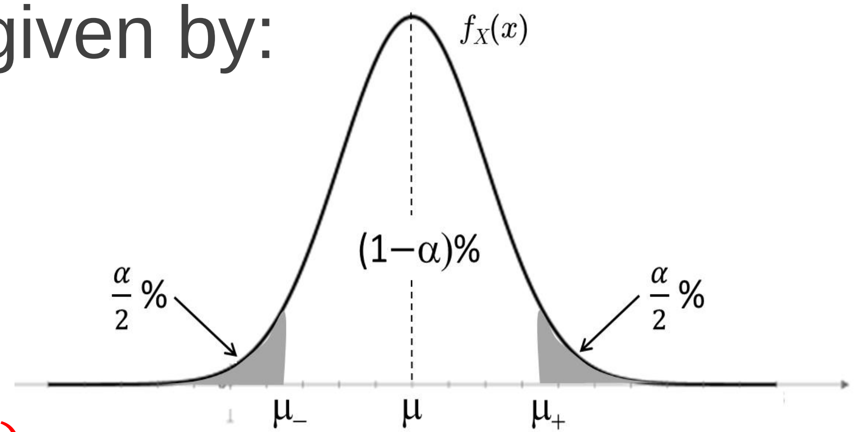
Confidence Interval

Conficence Interval

- The $1-\alpha$ confidence interval is an interval $[\theta_-;\theta_+]$ such that the probability that the true value of the unknown parameter θ lies within the interval is $1-\alpha$:

$$Pr(\theta_- \leq \theta \leq \theta_+) = 1-\alpha$$

- For a two-tailed normally distributed event with significance level α the $1-\alpha$ confidence interval for the mean $[\mu_-;\mu_+]$ is given by:



$$Pr(\mu_- \leq \mu \leq \mu_+) = 1-\alpha$$

- where the interval endpoints are:

$\Phi^{-1}(1-\frac{\alpha}{2})$

$$\mu_- = \bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \qquad \text{and} \qquad \mu_+ = \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

- For the 95% confidence interval: $z_{0.025} = \Phi^{-1}(0.975) = 1.96$

Cup Example

- Statistic model: $X \sim \mathcal{N}(250, 2.5^2)$
- Statistic: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- Estimate: $\hat{\mu} = \bar{X} = \frac{1}{25}\sum_{i=1}^{25} X_i = 250.2$
- Null hypothesis: $H_0 : \mu = 250$; Alternative hypothesis: $H_1 : \mu \neq 250$
- z-score: $z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{250.2-250}{2.5/\sqrt{25}} = 0.4$
- p-value $= Pr(Z > 0.4 \cup Z < -0.4) = 2 \cdot \left(1 - \Phi(0.4)\right) = 0.69 > 0.05 = \alpha$
- 95% confidence interval:

$$[\mu_-; \mu_+] = \left[\bar{x} - 1.96 \cdot \frac{\sigma}{\sqrt{n}};\ \bar{x} + 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right] = [249.22;\ 251.18]$$

ie. the Null hypothesis μ=250 can't be rejected

Tests and Types of Errors

- There will sometimes be wrong conclusions in hypothesis testing
- We can classify the errors in hypothesis testing as:

Table of error types		**Null hypothesis H_0**	
		True	**False**
Hypothesis test result	**Reject**	**Type I Error** (False negative: α)	Correct inference (True negative: 1-β)
	Fail to reject	Correct inference (True positive: 1-α)	**Type II Error** (False positive: β)

- $Pr(Type\ I\ Error) = Pr(Reject\ H_0 |\ H_0\ true) = \alpha$ (the significance level)
- $Pr(Type\ II\ Error) = Pr(Fail\ to\ reject\ H_0 |\ H_0\ false) = \beta$
- Decreasing the Type I Error rate (α) will increase the Type II Error rate (β)

Test catalog for the mean (known variance)

Statistical model:

- $X_1, X_2, \ldots X_n$ are i.i.d. samples of a random variable X with known mean μ (H_0) and variance σ^2.
- Parameter estimate:

  Sample mean: $\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$

- NOTE: The statistical model is only true if n is sufficiently large ($n > 30$) or if the samples are drawn from a normal population with mean μ and variance σ^2.

Hypothesis test (<u>two-tailed</u>):

- NULL hypothesis: $H_0\colon\ \mu = \mu_0$
- Alternative hypothesis: $H_1\colon\ \mu \neq \mu_0$
- Test size: $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$
- p-value: $pval = 2 \cdot (1 - \Phi(|z|))$

Confidence interval $[\mu_-, \mu_+]$:

- $z_{\alpha/2} = \Phi^{-1}(1 - \frac{\alpha}{2})$
- Lower bound: $\mu_- = \bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$
- Upper bound: $\mu_+ = \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

t-Score

When the variance is unknown!

Sample/empirical variance

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

- Now, consider the usual *z* statistic

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

- The equivalent statistic, when replacing the standard deviation (σ) with the empirical standard deviation (s), is called a *t*-score

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Called the student's t-distribution

- The *t*-score is *not* normally distributed; it is *t*-distributed with $\nu = n - 1$ degrees of freedom, which we write

$$\boxed{t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t(n-1)}$$

n-1 degrees of freedom

Student’s t-distribution

- **Students t-distribution: t(ν)=t(n-1):**

ν: Degrees of freedom

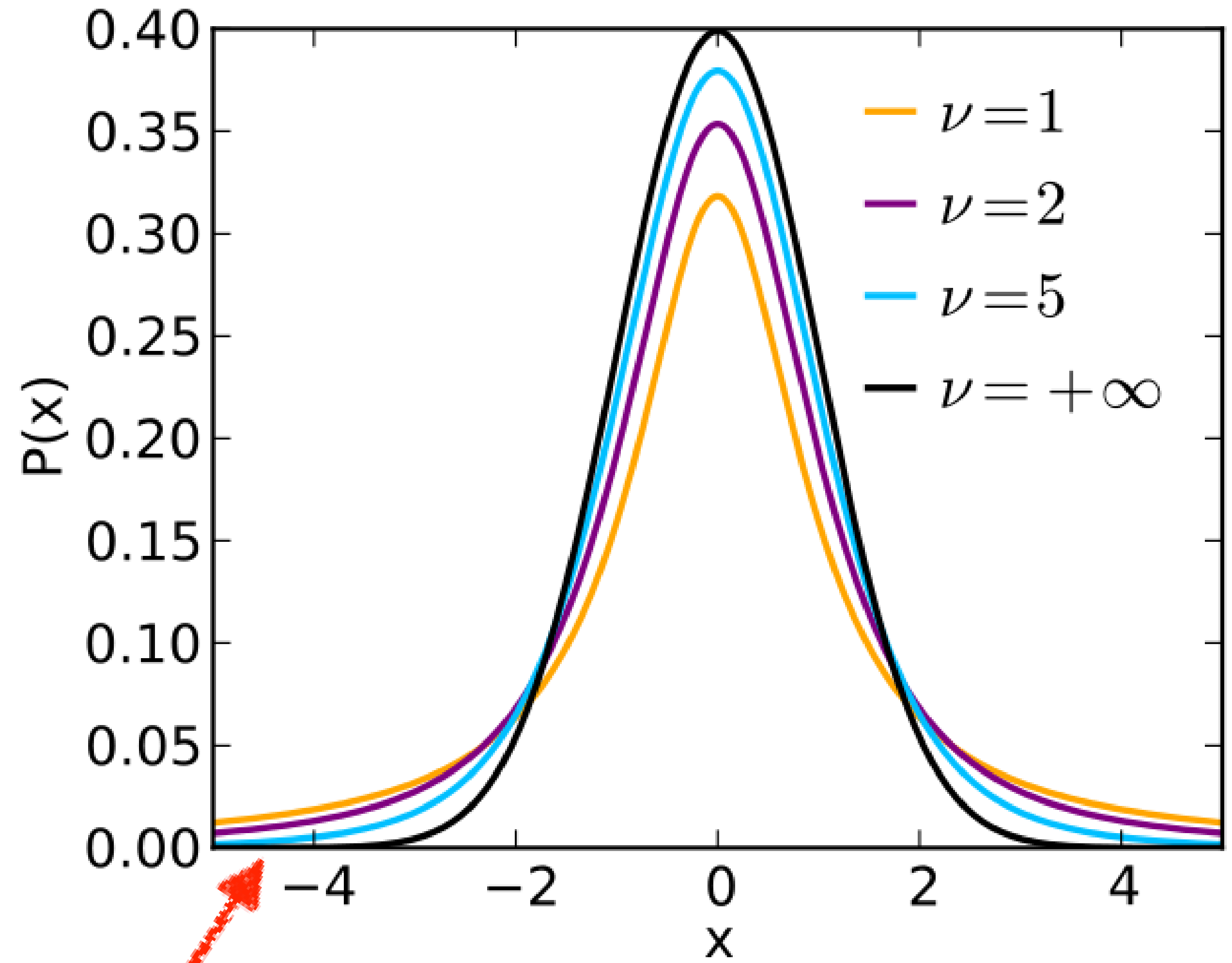
n: Number of samples

  - pdf: $f_X(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}(1+\frac{x^2}{\nu})^{-\frac{\nu+1}{2}}$

    where the gamma-function:

    $$\Gamma(n) = \int_0^\infty y^{n-1} e^{-y} dy$$

  - Even/symmetric: $f_X(x) = f_X(-x)$
  - Mean: μ_t = 0 for ν>1
  - Variance: $\sigma_t^2 = \frac{\nu}{\nu-2}$ for ν>2
  - $n \to \infty$: $t(n-1) \sim \mathcal{N}(0,1)$



t(ν) is heavy (large tail)
for small ν

When the variance is unknown!

Find the Mean Using the t-score

- To test the null hypothesis H_0: $\mu = \mu_0$ using the t-statistic instead of the z-statistic, the p-value is:

$$pval = 2 \cdot (1 - t_{cdf}(|t|, n-1))$$

- where $t_{cdf}(t, n-1) = Pr(T \leq t)$ denotes the CDF of a t distribution with n-1 degrees of freedom.

- The $(1 - \alpha)$% confidence interval for the mean is: $\bar{x} \pm t_{\alpha/2,\, n-1} \cdot s/\sqrt{n}$

  where $t_{\alpha/2,\, n-1} = t_0$ is chosen such that:

$$Pr\left(T \leq t_{\alpha/2,\, n-1}\right) = t_{cdf}\left(t_{\alpha/2,\, n-1}, n-1\right) = 1 - \frac{\alpha}{2}$$

- ie: $$t_0 = t_{\alpha/2,\, n-1} = t_{cdf,n-1}^{-1}\left(1 - \frac{\alpha}{2}\right)$$

Depends on n-1
(degrees of freedom)

The t-Distribution in Matlab (Mathcad)

- Calculating probabilities of a t-distributed random variable

$$T \sim t(n-1)$$

- $f(t)$ = `tpdf(t,n-1)` (`dt(t,n-1)`)
- $\Pr(T \leq t)$ = `tcdf(t,n-1)` (`pt(t,n-1)`)

- where n is the number of samples.
- Given a probability $1-\alpha/2$, what is the corresponding value t_0, such that $\Pr(T \leq t_0) = 1-\alpha/2$?

two-tailed event

- `t0 = tinv(1-alpha/2,n-1)` (`qt(1-α/2,n-1)`)

Cup Example

```
X =[253.50
    254.70
    256.23
    244.14
    252.78
    247.71
    249.52
    253.57
    248.86
    250.86
    251.87
    247.90
    249.59
    246.69
    251.45
    252.83
    249.50
    245.41
    250.19
    250.21
    250.17
    246.27
    248.24
    251.60
    251.21];
```

Data (n=25)

Then the sample mean ($\bar{X}$) is:

```
>>MeanX=mean(X)
            MeanX = 250.2000
```

And the (unbiased) estimate of the variance (s^2) is:

```
>>VarX=var(X)
            VarX = 8.5868
```

Corresponding to an empitical standard diviation (s):

```
>>StdX=sqrt(VarX)
            StdX = 2.9303
```

Cup Example

- Recall that in the cup filling machine example, the null hypothesis is:

$$H_0: \quad \mu = 250$$

- Test size:

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{250.20 - 250}{2.9303/\sqrt{25}} = 0.3413 \sim t(n-1)$$

- p-value:

$$pval = 2 \cdot \left(1 - t_{cdf}(|t|, n-1)\right) = 2 \cdot \left(1 - t_{cdf}(0.3413, 24)\right)$$

$$= 2 \cdot (1 - 0.6321) = 0.7359 \qquad > 0.05$$

- and we fail to reject the hypothesis

Cup Example

- The 95% confidence interval for the mean is $\bar{x} \pm t_0 \cdot \frac{s}{\sqrt{n}}$, so the endpoints are:

Lower bound: $$\mu_- = \bar{x} - t_0 \cdot \frac{s}{\sqrt{n}} = 250.20 - 2.0639 \cdot \frac{2.93}{\sqrt{25}} = 248.99$$

Upper bound: $$\mu_+ = \bar{x} + t_0 \cdot \frac{s}{\sqrt{n}} = 250.20 + 2.0639 \cdot \frac{2.93}{\sqrt{25}} = 251.41$$

- where

$$t_0 = tinv\left(1 - \frac{\alpha}{2}; n - 1\right) = tinv(0.975, 24) = 2.0639$$

TEST CATALOG FOR THE MEAN (UNKNOWN VARIANCE)

Statistical model:

- $X_1, X_2, \ldots X_n$ are i.i.d. samples of a random variable X with known mean μ (H_0) and unknown variance σ^2.
- Parameter estimate:

Sample mean: $\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$

Sample variance: $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

- NOTE: The statistical model is only true if n is sufficiently large ($n > 30$) or if the samples are drawn from a normal population with mean μ and variance σ^2.

Hypothesis test (two-tailed):

- NULL hypothesis: H_0: $\mu = \mu_0$
- Alternative hypothesis: H_1: $\mu \neq \mu_0$
- Test size: $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \sim t(n-1)$
- p-value: $pval = 2 \cdot (1 - t_{cdf}(|t|, n-1))$

Confidence interval [$\boldsymbol{\mu_-}$, $\boldsymbol{\mu_+}$]:

- $t_{\alpha/2,\, n-1} = {t_{cdf}}^{-1}(1 - \frac{\alpha}{2}, n-1)$ (often just called t_0)
- Lower bound: $\mu_- = \bar{x} - t_{\alpha/2,\, n-1} \cdot \frac{s}{\sqrt{n}}$
- Upper bound: $\mu_+ = \bar{x} + t_{\alpha/2,\, n-1} \cdot \frac{s}{\sqrt{n}}$

Convergence of the t-dist. towards a Std. Norm. dist.

More knowledges

- **z-test (Standard Normal distribution / known variance)):**
- p-value: $pval = 0.6892$
- Confidence interval: $[\mu_-; \mu_+] = [249.22; 251.18\,]$

Less knowledges

- **t-test (Student's t-distribution / unknown variance):**
- p-value: $pval = 0.7359$ p-value larger → More difficult to reject H_0
- Confidence interval: $[\mu_-; \mu_+] = [248.99; 251.41\,]$

Confidence interval wider → True value more uncertain

Convergence of the t-dist. towards a Std. Norm. dist.

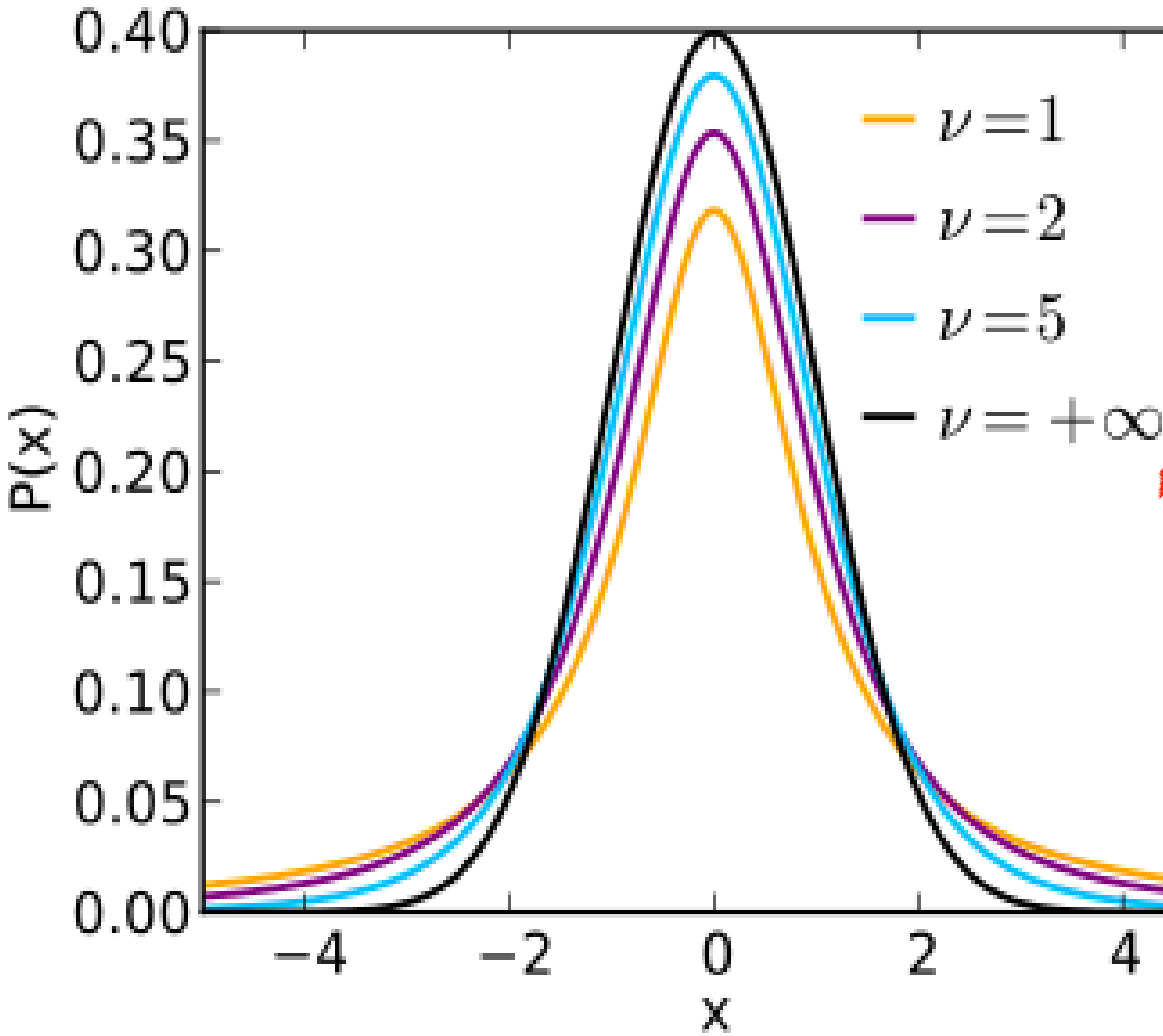
tinv(0.975, ν)

Degrees of freedom

ν	t_0
2	12,71
3	4,30
5	2,78
10	2,26
30	2,05
∞	1,96

$\phi^{-1}(0.975)$

density of the t-distribution



Standard normal distribution

Checking for Normality in Sampled Data (Q-Q plots)

- **Large sample size (say $n \geq 30$):**
  - ➢ The central limit theorem (CLT) quite safely allow us to make interference about the mean of any population (i.e. distribution) using z-score or t-score.

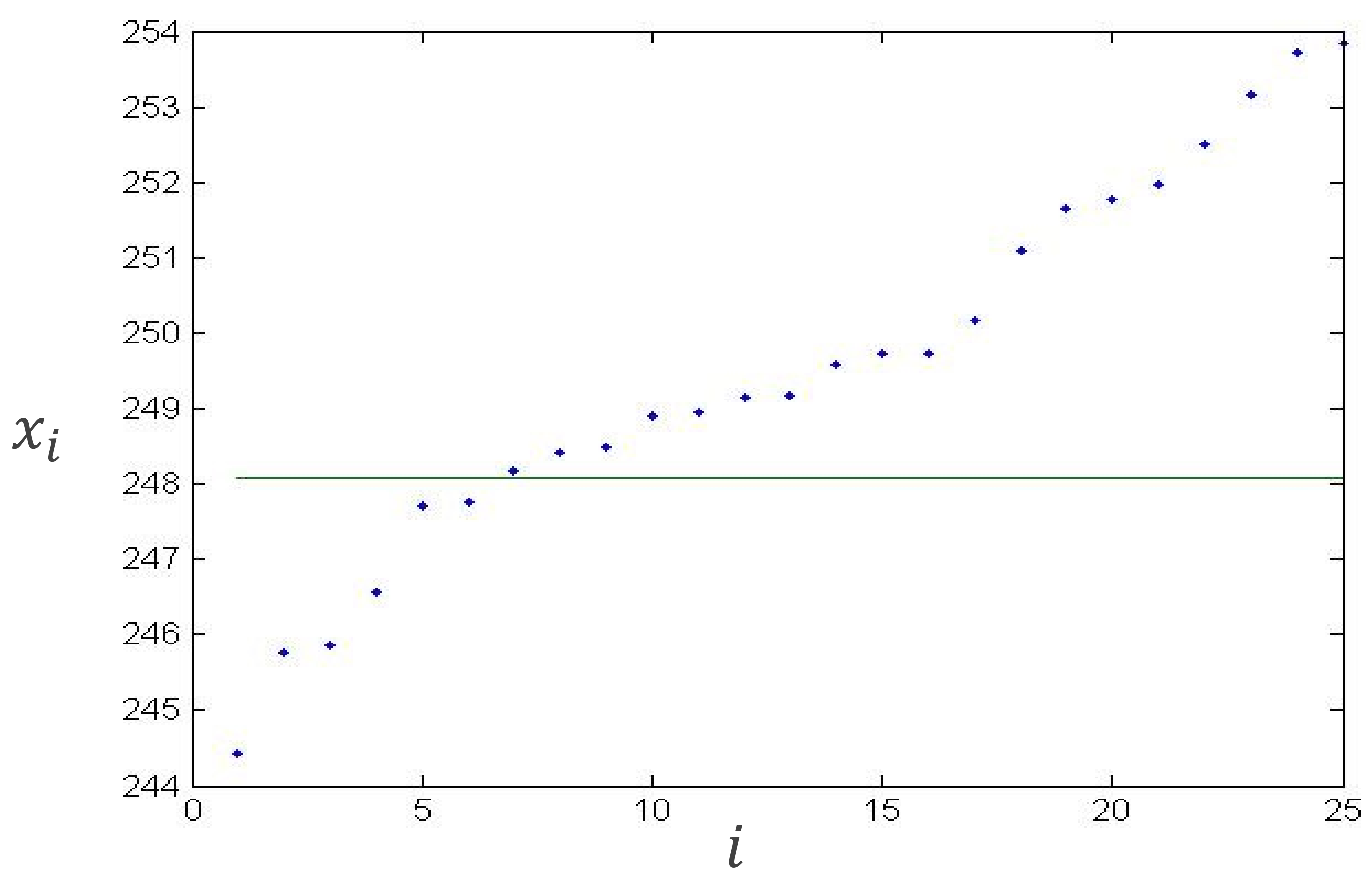
- **Small sample size (say $n < 30$):**
  - ➢ CLT does not hold anymore.
  - ➢ Statistical interference based on either z-score or t-score only works, if the sampled data $x_1, x_2, \ldots, x_n$ are themselves normally distributed.
  - ➢ **Hence, we need a method to check whether the data are normally distributed!**

Quantiles *(Fraktiler)*

- The 25% quantile of the previous data

```
q25 = quantile(x,0.25)
q25 = 248.0731
```

$$\Pr(X \le q25) = \Phi(z25) = 0.25$$
$$\Downarrow$$
$$z25 = \frac{q25 - \mu}{\sigma/\sqrt{n}} = \Phi^{-1}(0.25) = -0.675$$



Sorted data values with the estimated 25% percentile = 248.07.

Roughly 25% of the data should lie below this value.

Q-Q plot

- The quantiles of standard normally distributed data with n samples are roughly such that:

$$\sim[0;1]$$

$$\boldsymbol{x}_{[i]} \leftrightarrow \Phi_{[i]}^{-1}\left(\frac{\boldsymbol{i}-\mathbf{0.5}}{\boldsymbol{n}}\right)=\boldsymbol{z}_{[i]}=\frac{\boldsymbol{x}_{[i]}-\boldsymbol{\mu}}{\boldsymbol{\sigma}} \longrightarrow \boldsymbol{x}_{[i]}=\boldsymbol{\sigma}\cdot\Phi_{[i]}^{-1}+\boldsymbol{\mu}$$

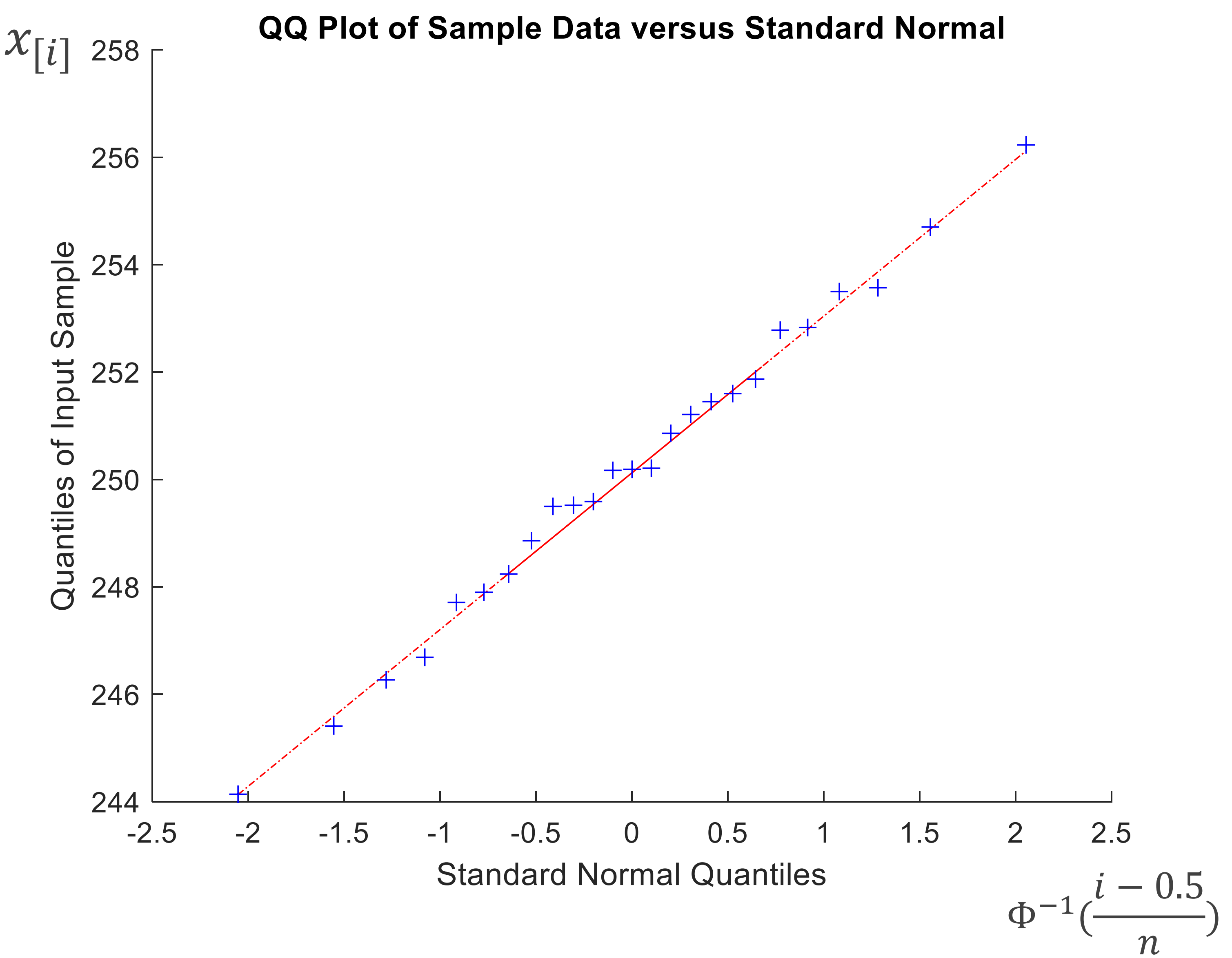
$$]-\infty;\infty[$$

- where $\boldsymbol{x}_{[i]}$ denotes the $\boldsymbol{i}$'th sample after sorting the samples $x_1, x_2, \ldots, x_n$ in ascending order.
- If the data are consistent with a sample from a normal distribution then plotting $\boldsymbol{x}_{[i]}$ vs. $\Phi_{[i]}^{-1}\left(\frac{\boldsymbol{i}-\mathbf{0.5}}{\boldsymbol{n}}\right)$ should result in a straight line.
- **This is the Q-Q plot.**

Q-Q plot

- **Q-Q plot – a method to check whether the data are normally distributed**

- Sort the samples in ascending order: $x_1, x_2, \ldots, x_n$
- Plot $x_{[i]}$ vs. $\Phi^{-1}(\frac{i-0.5}{n})$
- If the data are consistent with a sample from a normal distribution it should result in a straight line.
- In Matlab/Mathcad: `qqplot(x)`

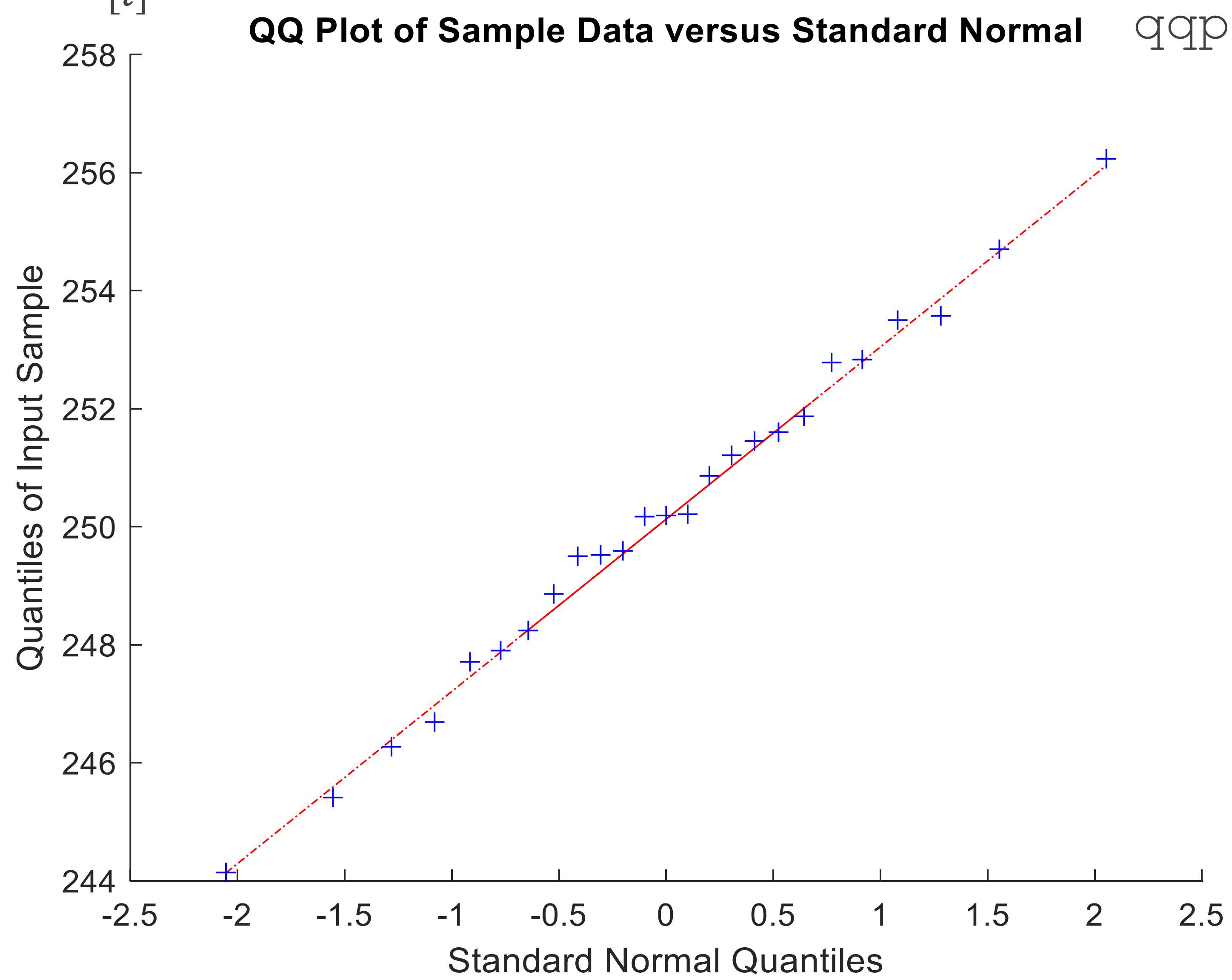


Cup Example

```
X =[253.50
    254.70
    256.23
    244.14
    252.78
    247.71
    249.52
    253.57
    248.86
    250.86
    251.87
    247.90
    249.59
    246.69
    251.45
    252.83
    249.50
    245.41
    250.19
    250.21
    250.17
    246.27
    248.24
    251.60
    251.21];
```

$x_{[i]}$

`qqplot(x)`



$$\Phi^{-1}\left(\frac{i-0.5}{n}\right)$$

Q-Q plot of the cup-data. The data points lie roughly on a straight line, and we conclude that the data are in fact normally distributed.

Words and Concepts to Know

Heavy

Null hypothesis

Left-tailed

Test catalog

Type I Error

Reject

Alternative hypothesis

t-score

Type II Error

Right-tailed

Fail to reject

Students t-distribution

Quantiles

Q-Q plot

Two-tailed

Degrees of freedom

Hypothesis test