# Chapter II: Basics from Linear Algebra, Probability Theory, and Statistics

Information Retrieval & Data Mining
Universität des Saarlandes, Saarbrücken
Wintersemester 2013/14

# Chapter II

**II.1 Linear Algebra**
Vectors, Matrices, Eigenvalues, Eigenvectors,
Singular Value Decomposition

**II.2 Probability Theory**
Events, Probabilities, Random Variables, Distributions,
Bounds, Limit Theorems

**II.3 Statistical Inference**
Parameter Estimation, Confidence Intervals, Hypothesis Testing

# II.3 Statistical Inference

1. **Parameter Estimation**
2. **Confidence Intervals**
3. **Hypothesis Testing**

Based on **LW Chapters 6, 7, 9, 10**

# Statistical Model

- A **statistical model** $M$ is a set of distributions (or regression functions), e.g., all unimodal smooth distributions
- $M$ is called a **parametric model** if it can be completely described by a finite number of parameters, e.g., the family of Normal distributions for a finite number of parameters $\mu$ and $\sigma$

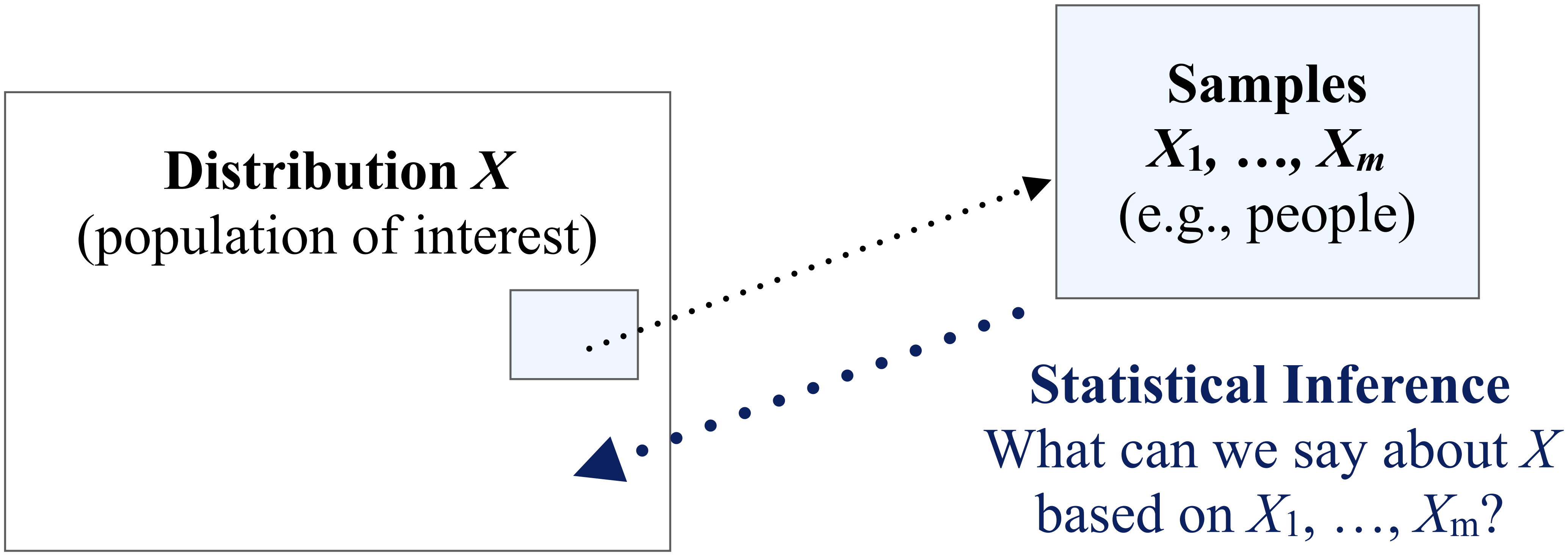
$$M = \left\{ f_X(x; \mu, \sigma) = \frac{1}{\sqrt{2\,\pi\,\sigma}}\, e^{-\frac{(x-\mu)^2}{2\,\sigma^2}} \;\middle|\; \mu \in \mathbb{R},\ \ \sigma > 0 \right\}$$

# Statistical Inference

- Given a parametric model $M$ and a **sample** $X_1, \ldots, X_m$, how do we infer (learn) the parameters of $M$?

- For multivariate models with observed variable $X$ and response variable $Y$, this is called **prediction** or **regression**, for a discrete outcome variable this is also called **classification**

# Idea of Sampling




- Example: Suppose we want to estimate the average salary of employees in German companies
  - Sample 1: Suppose we look at $n = 200$ top-paid CEOs of major banks
  - Sample 2: Suppose we look at $n = 1{,}000$ employees across all sectors

# Basic Types of Statistical Inference

- Given **independent and identically distributed** (iid.) samples $X_1, \dots, X_n \sim X$ of an unknown distribution $X$
  - e.g.: $n$ single-coin-toss experiments $X_1, \dots, X_n \sim \text{Bernoulli}(p)$
- **Parameter estimation**
  - e.g.: what is the parameter $p$ of Bernoulli($p$)?
    what is $E[X]$, the cdf $F_X$ of $X$, the pdf $f_X$ of $X$, etc.?
- **Confidence intervals**
  - e.g.: give me all values $C = [a, b]$ such that $P[p \in C] \geq 0.95$
    with interval boundaries $a$ and $b$ derived from samples $X_1, \dots, X_n$
- **Hypothesis testing**
  - e.g.: $H_0 : p = 1/2$ (i.e., coin is fair) vs. $H_1 : p \neq 1/2$

# 1. Parameter Estimation

- A **point estimator** for a parameter $\theta$ of a probability distribution $X$ is a random variable $\hat{\theta}_n$ derived from an iid. sample $X_1, \ldots, X_n$

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

- An estimator $\hat{\theta}_n$ for parameter $\theta$ is **unbiased** if $E[\hat{\theta}_n] = \theta$

  otherwise the estimator has **bias** $E[\hat{\theta}_n] - \theta$

- An estimator on sample size $n$ is **consistent** if

$$\lim_{n\to\infty} P[|\hat{\theta}_n - \theta| < \epsilon] = 1 \text{ for any } \epsilon > 0$$

# Estimation Error

- Let $\hat{\theta}_n$ be an estimator for parameter $\theta$ over iid. samples $X_1, \ldots, X_n$
- The distribution of $\hat{\theta}_n$ is called **sampling distribution**
- The **standard error** for $\hat{\theta}_n$ is: $se(\hat{\theta}) = \sqrt{Var(\hat{\theta}_n)}$
- The **mean squared error** (MSE) for $\hat{\theta}_n$ is:

$$MSE(\hat{\theta}_n) = E[(\hat{\theta}_n - \theta)^2] = bias^2(\hat{\theta}_n) + Var(\hat{\theta}_n)$$

- The estimator $\hat{\theta}_n$ is **asymptotically Normal** if

$(\hat{\theta}_n - \theta)/se$ converges in distribution to N(0,1)

# Types of Estimation

- **Non-Parametric Estimation**
  - **no assumptions** about the model $M$ nor the parameters $\theta$ of the underlying distribution $X$
  - e.g.: "plug-in estimators" (e.g., histograms) to approximate $X$
- **Parametric Estimation**
  - **requires assumptions** about the model $M$ and the parameters $\theta$ of the underlying distribution $X$
  - analytical or numerical methods for estimating $\theta$
    - **Method of Moments**
    - **Maximum Likelihood**
    - **Expectation Maximization** (EM)

# Empirical Distribution Function

- The **empirical distribution function** $\hat{F}_n$ is the cdf that puts probability mass $1/n$ at each data point $X_i$

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(X_i \leq x)$$

with indicator function

$$\mathbb{I}(X_i \leq x) = \begin{cases} 1 & : \quad X_i \leq x \\ 0 & : \quad X_i > x \end{cases}$$

- A **statistical function** ("statistics") $T(F)$ is any function over $F$, e.g., mean, variance, skewness, median, quantiles, correlation
- The **plug-in estimator** of $\theta = T(F)$ is $\hat{\theta}_n = T(\hat{F}_n)$

# Histograms as Density Estimators

- Instead of the full empirical distribution, often compact synopses can be used, such as **histograms** where $X_1, \ldots, X_n$ are grouped into $m$ **cells** (buckets) $c_1, \ldots, c_m$ with bucket boundaries $lb(c_i)$ and $ub(c_i)$

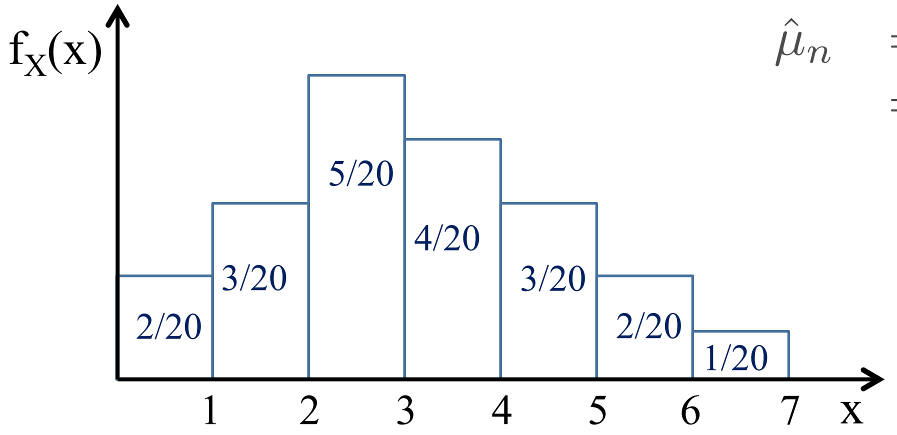
$$lb(c_1) = -\infty, \quad ub(c_m) = \infty, \quad ub(c_{i-1}) = lb(c_i) \text{ for } (1 \leq i \leq m), \quad \text{and}$$

$$freq_f(c_i) = \hat{f}_n(x) = \frac{1}{n} \sum_{j=1}^{n} \mathbb{I}(lb(c_i) < X_j \leq ub(c_i))$$

$$freq_F(c_i) = \hat{F}_n(x) = \frac{1}{n} \sum_{j=1}^{n} \mathbb{I}(X_j \leq ub(c_i))$$

- Example:

$X_1 = X_2 = 1$
$X_3 = X_4 = X_5 = 2$
$X_6 = \ldots X_{10} = 3$
$X_{11} = \ldots X_{14} = 4$
$X_{15} = \ldots X_{17} = 5$
$X_{18} = X_{19} = 6$
$X_{20} = 7$




$$\begin{aligned}\hat{\mu}_n &= 1 \times \tfrac{2}{20} + 2 \times \tfrac{3}{20} + \ldots + 7 \times \tfrac{1}{20} \\ &= 3.65\end{aligned}$$

# Method of Moments

- Suppose parameter $\theta = (\theta_1, \ldots, \theta_k)$ has $k$ components
- Compute ***j*-th moment** for $1 \le j \le k$:

$$\alpha_j = \alpha_j(\theta) = E_\theta[X^j] = \int_{-\infty}^{+\infty} x^j \, f_X(x) \, dx$$

- Compute ***j*-th sample moment** for $1 \le j \le k$:

$$\hat{\alpha}_j = \frac{1}{n} \sum_{i=1}^{n} X_i^j$$

- **Method-of-moments** estimate of $\theta$ is obtained by solving a system of $k$ equations in $k$ unknowns

$$\begin{gathered} \alpha_1(\hat{\theta}_n) = \hat{\alpha}_1 \\ \vdots \\ \alpha_k(\hat{\theta}_n) = \hat{\alpha}_k \end{gathered}$$

# Method of Moments (Example)

- Let $X_1, \ldots, X_n \sim \text{Normal}(\mu, \sigma^2)$.

$$\alpha_1 = E_\theta[X] = \mu$$

$$\alpha_2 = E_\theta[X^2] = Var(X) + (E[X])^2 = \sigma^2 + \mu^2$$

- By solving the **system of 2 equations** in 2 unknowns

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$\hat{\sigma}^2 + \hat{\mu}^2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2$$

we obtain as solutions

$$\hat{\mu} = \bar{X}_n \qquad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$$

# Maximum Likelihood

- Let $X_1, \ldots, X_n$ be iid. with pdf $f(x;\theta)$
- Estimate parameter $\theta$ of a **postulated distribution** $f(x;\theta)$ such that the likelihood that the sample values $x_1, \ldots, x_n$ are generated by the distribution are maximized
- Maximize $L(x_1, \ldots, x_n, \theta) \approx \mathrm{P}[x_1, \ldots, x_n$ originate from $f(x;\theta)]$
- Usually formulated as:

$$\underset{\theta}{arg\,max}\ L_n[\theta] = \prod_{i=1}^{n} f(X_i, \theta)$$

- The value $\hat{\theta}$ that maximizes $L_n[\theta]$ is called the **maximum-likelihood estimate (MLE)** of $\theta$
- If analytically intractable, MLE can be determined using numerical iteration methods

# Maximum Likelihood (Example)

- Let $X_1, \ldots, X_n \sim$ Bernoulli($p$) (corresponding to $n$ coin tosses)
- Assume that we observed $h$ times head and ($n$-$h$) times tail
- **Maximum-likelihood estimation** of parameter $p$

$$L[h,n,p] = \prod_{i=1}^{n} f(X_i;p) = \prod_{i=1}^{n} p^{X_i}(1-p)^{1-X_i} = p^h\,(1-p)^{(n-h)}$$

- Maximize **log-likelihood function**

$$log\,L[h,n,p] = h \times log(p) + (n-h) \times log(1-p)$$

$$\frac{\partial L}{\partial p} = \frac{h}{p} - \frac{n-h}{1-p} = 0 \quad \Rightarrow \quad p = \frac{h}{n}$$

# Maximum Likelihood for Normal Distributions

$$L(x_1, \ldots, x_n, \mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \prod_{i=1}^{n} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$\frac{\partial L}{\partial \mu} = \frac{-1}{2\sigma^2} \sum_{i=1}^{n} 2(x_i - \sigma) = 0$$

$$\frac{\partial L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2 = 0$$

$$\Rightarrow \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu})^2$$

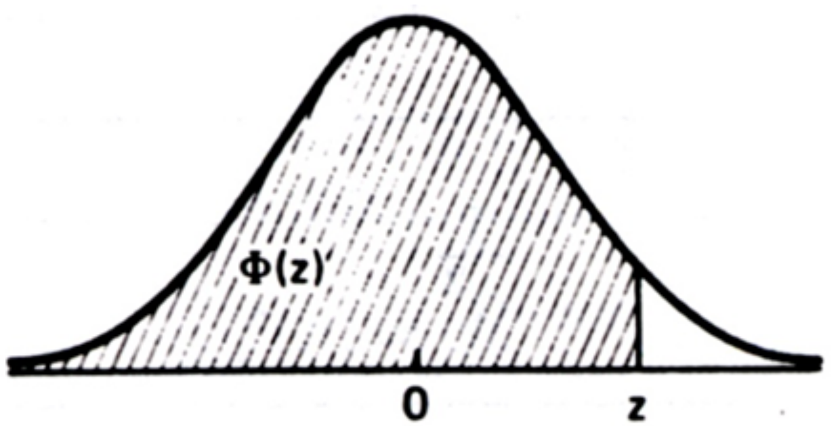
# 2. Confidence Intervals

- Determine **interval estimator** $T$ for parameter $\theta$ such that

$$P[T - a \leq \theta \leq T + a] = 1 - \alpha$$

$T \pm a$ is the **confidence interval** and $1\text{-}\alpha$ the **confidence level**

- For the distribution of a random variable $X$, a value $x_\gamma$ $(0 < \gamma < 1)$ is with $P[X \leq x_\gamma] \geq \gamma$ and $P[X \geq x_\gamma] \geq 1\text{-}\gamma$ is called **$\gamma$-quantile**
  - the 0.5-quantile is known as **median**
  - for the standard Normal distribution N(0,1) the $\gamma$-quantile is denoted $\Phi_\gamma$
- For a given a <u>or</u> $\alpha$, find a value $z$ of N(0,1) that denotes the [T-a,T+a] confidence interval <u>or</u> a corresponding $\gamma$-quantile for $1\text{-}\alpha$

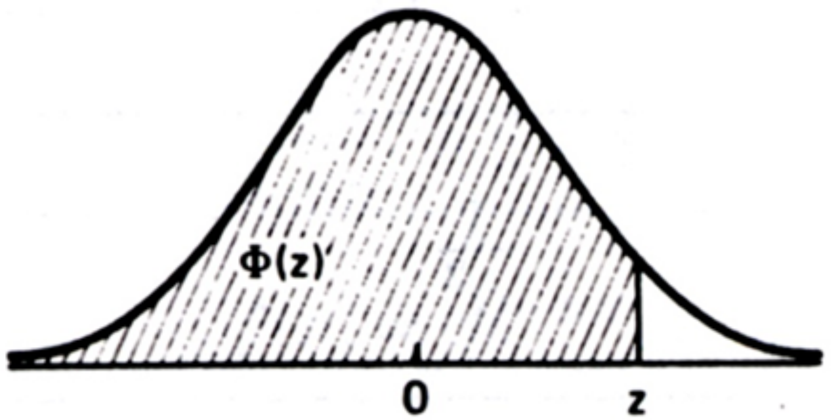


# Confidence Intervals for Expectations (I)

- Let $X_1, \ldots, X_n$ be a sample from a distribution $X$ with unknown expectation $\mu$ and known variance $\sigma^2$
- For sufficiently large $n$, the sample mean $\bar{X}$ is $N(\mu, \sigma^2/n)$ distributed and

$$
\begin{aligned}
P[-z \leq \tfrac{(\bar{X}-\mu)\sqrt{n}}{\sigma} \leq z] &= \Phi(z) - \Phi(-z) \\
&= \Phi(z) - (1 - \Phi(z)) \\
&= 2\,\Phi(z) - 1 \\
&= P[\bar{X} - \tfrac{z\,\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \tfrac{z\,\sigma}{\sqrt{n}}]
\end{aligned}
$$

$$
\Rightarrow \quad P[\bar{X} - \frac{\Phi_{1-\alpha/2}\,\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{\Phi_{1-\alpha/2}\,\sigma}{\sqrt{n}}] = 1 - \alpha
$$



# Confidence Intervals for Expectations (I) (cont'd)

$$P[\bar{X} - \frac{\Phi_{1-\alpha/2}\,\sigma}{\sqrt{n}} \le \mu \le \bar{X} + \frac{\Phi_{1-\alpha/2}\,\sigma}{\sqrt{n}}] = 1 - \alpha$$

- For **confidence interval** $[\bar{X} - a, \bar{X} + a]$ compute

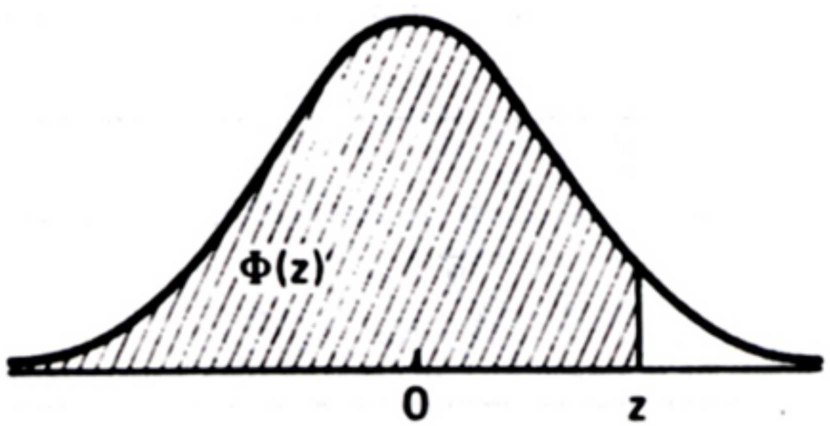
  $z = \frac{a\sqrt{n}}{\sigma}$ and lookup Φ(z) to **determine 1-*α***

- For **confidence level** 1-*α* set

  $z = \Phi_{1-\frac{\alpha}{2}}$ (i.e., as (1-*α*/2)-quantile of N(0,1))

  then $a = \frac{z\,\sigma}{\sqrt{n}}$ to determine

  confidence interval



# Confidence Intervals for Expectations (I) (Example)

- Based on a random sample of $n = 100$ queries, we observe an average response time of $\bar{X} = 64$ . We further know that the standard deviation is $\sigma = 4$

- Q: What is the confidence of the interval 64±0.5?

$$\begin{aligned} a &= 0.5 \\ z &= \frac{0.5\sqrt{100}}{4} = 1.25 \\ \Phi(1.25) &= 0.89435 \\ 1 - \frac{\alpha}{2} &= 0.89435 \\ 1 - \alpha &= 0.7887 \end{aligned}$$

A: 78.87%

- Q: What's the 99% confidence interval?

$$\begin{aligned} 1 - \alpha &= 0.99 \\ \alpha &= 0.01 \\ a &= \frac{\Phi_{0.005} \times 4}{\sqrt{100}} = 1.032 \end{aligned}$$

A: 64±1.032

# Confidence Intervals for Expectations (II)

- Let $X_1, \ldots, X_n$ be an iid. sample from a distribution $X$ with unknown expectation $\mu$, unknown variance $\sigma^2$, but known sample variance $S^2$
- For sufficiently large $n$, the random variable

$$T = \frac{(\bar{X} - \mu)\sqrt{n}}{S}$$

has a **Student's *t* distribution** with ($n$-1) degrees of freedom

$$f_{T,n}(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\,\pi}\left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}}$$

with the Gamma function

$$\Gamma(x) = \int_0^\infty e^{-t}\, t^{x-1}\, dt \quad \text{for} \quad x > 0$$

# Confidence Intervals for Expectations (II) (cont'd)

$$P[\bar{X} - \frac{t_{n-1,1-\alpha/2}\, S}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{t_{n-1,1-\alpha/2}\, S}{\sqrt{n}}] = 1 - \alpha$$

- For **confidence interval** $[\bar{X} - a, \bar{X} + a]$ compute

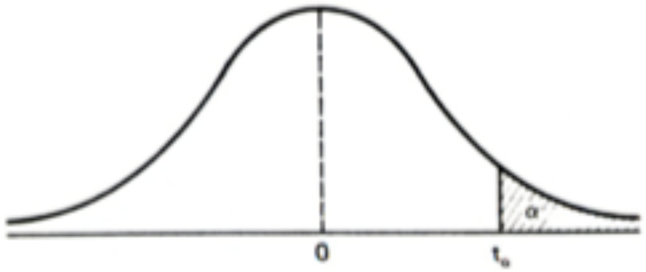
  $t = \frac{a\sqrt{n}}{S}$ and lookup $f_{T(n\text{-}1)}(t)$ to **determine 1-$\alpha$**

- For **confidence level** 1-$\alpha$ set

  $t = t_{n-1,1-\alpha/2}$ (i.e., as (1-$\alpha$/2)-quantile of $f_{T(n\text{-}1)}$)

  then $a = \frac{t\, S}{\sqrt{n}}$ to determine

  confidence interval



# 3. Hypothesis Testing

- Suppose we throw a coin $n$ times and want to know whether the coin is fair, i.e., $P(H) = P(T)$
- Let $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$ be the iid. coin flips, so that the coin is fair if $p = 0.5$
- Let the **null hypothesis** $H_0$ be “the coin is fair”
- The **alternative hypothesis** $H_1$ is then “the coin is not fair”
- Intuitively, if $|\bar{X} - 0.5|$ is large, we should **reject** $H_0$

# Hypothesis Testing Terminology

- $\theta = \theta_0$ is called a **simple hypothesis**
- $\theta > \theta_0$ or $\theta < \theta_0$ is called a **compound hypothesis**
- $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$ is called a **two-sided test**
- $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$ and $H_0 : \theta \geq \theta_0$ vs. $H_1 : \theta < \theta_0$ are called a **one-sided test**
- **Rejection region** $R$ : if $X \in R$, reject $H_0$ otherwise retain $H_0$
- The rejection region is typically defined using a **test statistic** $T$ and a **critical value** $c$

$$R = \{ X \ : \ T(X) > c \}$$

# $p$-Values

- The $p$-value is the probability that **if $H_0$ holds**, we observe values **at least as extreme** of the test statistic
  - It is not the probability that $H_0$ holds
  - The **smaller** the $p$-value, the **stronger** is the evidence against $H_0$, i.e., if we observe a small enough $p$-value, we can reject $H_0$
  - How small the $p$-value needs to be depends on the application
- Typical $p$-value scale:
  - $< 0.01$ **very strong** evidence against $H_0$
  - $0.01 – 0.05$ **strong** evidence against $H_0$
  - $0.05 – 0.10$ **weak** evidence against $H_0$
  - $> 0.1$ **little or no** evidence against $H_0$

# Types of Errors & Statistical Significance

| | **Retain $H_0$** | **Reject $H_0$** |
|---|---|---|
| **$H_0$ true** | OK | **Type I Error** |
| **$H_1$ true** | **Type II Error** | OK |

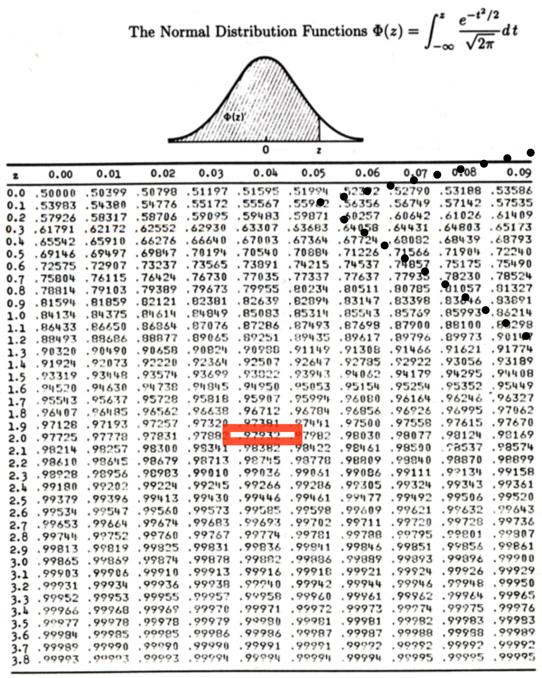
- Hypothesis tests often performed at a **level of significance α**
  - means that $H_0$ is rejected if the *p*-value is less than α
  - reported as "*results is **statistically significant at the α level***"
  - specifying *p*-values is more informative
- Don't confuse **statistical significance** with **practical significance**
  - e.g.: "*blue hyperlinks increase click rate by 0.0001% over black ones*"
    "*fuel consumption is reduced by 0.0001 l/km by new part*"
    …

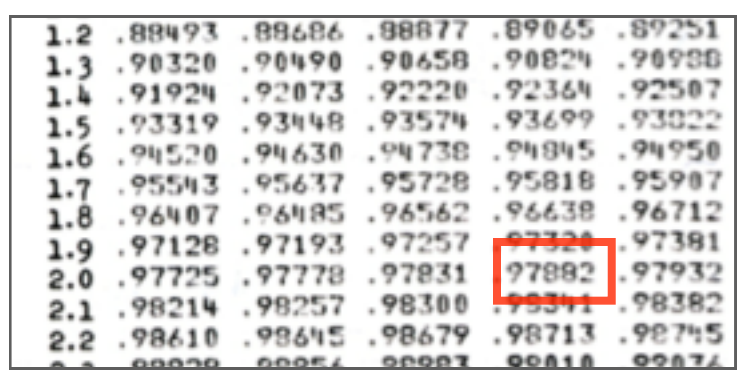
# The Wald Test

- Two-sided test for $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$
- Test statistic $W = \dfrac{|\hat{\theta} - \theta_0|}{\hat{se}}$ with sample estimate $\hat{\theta}$
  and $\hat{se} = se(\hat{\theta}) = \sqrt{Var(\hat{\theta})}$
- *W* converges in probability to N(0, 1)
- If *w* is the observed value of the Wald statistic,
  the *p*-value is $2\Phi(-|w|)$

# The Wald Test (Example)

- We can use the Wald test to test if our coin is fair
  - Suppose the **observed sample mean** is 0.6 and the **observed standard error** is 0.049
  - We obtain as a **test statistic value** $w = (0.6 - 0.5) / 0.049 \approx 2.04$
  - The ***p*-value** is therefore $2\Phi(-|2.04|) \approx 0.042$ (i.e., a fair coin would lead to such an extreme value $w$ only with probability 0.042), which gives us **strong evidence** to **reject** the null hypothesis $H_0$



| | | | | | |
|---|---|---|---|---|---|
| 1.2 | .88493 | .88686 | .88877 | .89065 | .89251 |
| 1.3 | .90320 | .90490 | .90658 | .90824 | .90988 |
| 1.4 | .91924 | .92073 | .92220 | .92364 | .92507 |
| 1.5 | .93319 | .93448 | .93574 | .93699 | .93822 |
| 1.6 | .94520 | .94630 | .94738 | .94845 | .94950 |
| 1.7 | .95543 | .95637 | .95728 | .95818 | .95907 |
| 1.8 | .96407 | .96485 | .96562 | .96638 | .96712 |
| 1.9 | .97128 | .97193 | .97257 | .97320 | .97381 |
| 2.0 | .97725 | .97778 | .97831 | $\boxed{.97882}$ | .97932 |
| 2.1 | .98214 | .98257 | .98300 | .98341 | .98382 |
| 2.2 | .98610 | .98645 | .98679 | .98713 | .98745 |

$2 * (1 - 0.97882) \approx 0.042$

# Pearson's $\chi^2$ Test for Multinomial Data

- Let $X_1, \ldots, X_m \sim \text{Multinomial}(n, \mathbf{p})$, the MLE of $\mathbf{p}$ is $(X_1/n, X_2/n, \ldots, X_n/n)$
- Let $\mathbf{p_0} = (p_{01}, p_{02}, \ldots, p_{0n})$ and we want to test $\mathrm{H}_0 : \mathbf{p} = \mathbf{p_0}$ vs. $\mathrm{H}_1 : \mathbf{p} \neq \mathbf{p_0}$
- Pearson's $\chi^2$ statistic is

$$T = \sum_{j=1}^{k} \frac{(X_j - n\,p_{0j})^2}{n\,p_{0j}} = \sum_{j=1}^{k} \frac{(X_j - E_j)^2}{E_j}$$

with expected value $E_j = \mathrm{E}[X_j] = n\, p_{0j}$ of $X_j$ under $\mathrm{H}_0$

- The $p$-value is $P(\chi^2_{k-1} > t)$ where $t$ is the observed value of the test statistic and there are ($k$-1) degrees of freedom

# Pearson's $\chi^2$ Test for Multinomial Data (Example)

- We can use Pearson's $\chi^2$ test to test whether a **dice is fair**
  - Suppose after 1,000 throws of the dice, we observed
    ① x 173, ② x 167, ③ x 167, ④ x 176, ⑤ x 167, ⑥ x 150
    => **p** = (0.173, 0.167, 0.167, 0.176, 0.167, 0.150) (based on MLE)
  - $\mathbf{p_0}$ = (0.167, 0.167, 0.167, 0.167, 0.167, 0.167)
  - $T$ = 2.43 => $p$-value is 0.80 giving us no evidence to reject $H_0$

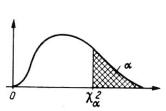



Abb. 1.4

| Anzahl der Freiheits-grade $m$ | Wahrscheinlichkeit $p = \alpha$ | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0,99 | 0,98 | 0,95 | 0,90 | 0,80 | 0,70 | 0,50 | 0,30 | 0,20 | 0,10 | 0,05 | 0,02 | 0,01 | 0,005 | 0,002 | 0,001 |
| 1 | 0,00016 | 0,0006 | 0,0039 | 0,016 | 0,064 | 0,148 | 0,455 | 1,07 | 1,64 | 2,7 | 3,8 | 5,4 | 6,6 | 7,9 | 9,5 | 10,83 |
| 2 | 0,020 | 0,040 | 0,103 | 0,211 | 0,446 | 0,713 | 1,386 | 2,41 | 3,22 | 4,6 | 6,0 | 7,8 | 9,2 | 10,6 | 12,4 | 13,8 |
| 3 | 0,115 | 0,185 | 0,352 | 0,584 | 1,005 | 1,424 | 2,366 | 3,67 | 4,64 | 6,3 | 7,8 | 9,8 | 11,3 | 12,8 | 14,8 | 16,3 |
| 4 | 0,30 | 0,43 | 0,71 | 1,06 | 1,65 | 2,19 | 3,36 | 4,9 | 6,0 | 7,8 | 9,5 | 11,7 | 13,3 | 14,9 | 16,9 | 18,5 |
| 5 | 0,55 | 0,75 | 1,14 | 1,61 | $\boxed{2,34}$ | 3,00 | 4,35 | 6,1 | 7,3 | 9,2 | 11,1 | 13,4 | 15,1 | 16,8 | 18,9 | 20,5 |
| 6 | 0,87 | 1,13 | 1,63 | 2,20 | 3,07 | 3,83 | 5,35 | 7,2 | 8,6 | 10,6 | 12,6 | 15,0 | 16,8 | 18,5 | 20,7 | 22,5 |
| 7 | 1,24 | 1,56 | 2,17 | 2,83 | 3,82 | 4,67 | 6,35 | 8,4 | 9,8 | 12,0 | 14,1 | 16,6 | 18,5 | 20,3 | 22,6 | 24,3 |
| 8 | 1,65 | 2,03 | 2,73 | 3,49 | 4,59 | 5,53 | 7,34 | 9,5 | 11,0 | 13,4 | 15,5 | 18,2 | 20,1 | 22,0 | 24,3 | 26,1 |
| 9 | 2,09 | 2,53 | 3,32 | 4,17 | 5,38 | 6,39 | 8,34 | 10,7 | 12,2 | 14,7 | 16,9 | 19,7 | 21,7 | 23,6 | 26,1 | 27,9 |
| 10 | 2,56 | 3,06 | 3,94 | 4,86 | 6,18 | 7,27 | 9,34 | 11,8 | 13,4 | 16,0 | 18,3 | 21,2 | 23,2 | 25,2 | 27,7 | 29,6 |
| 11 | 3,1 | 3,6 | 4,6 | 5,6 | 7,0 | 8,1 | 10,3 | 12,9 | 14,6 | 17,3 | 19,7 | 22,6 | 24,7 | 26,8 | 29,4 | 31,3 |
| 12 | 3,6 | 4,2 | 5,2 | 6,3 | 7,8 | 9,0 | 11,3 | 14,0 | 15,8 | 18,5 | 21,0 | 24,1 | 26,2 | 28,3 | 30,9 | 32,9 |
| 13 | 4,1 | 4,8 | 5,9 | 7,0 | 8,6 | 9,9 | 12,3 | 15,1 | 17,0 | 19,8 | 22,4 | 25,5 | 27,7 | 29,8 | 32,5 | 34,5 |
| 14 | 4,7 | 5,4 | 6,6 | 7,8 | 9,5 | 10,8 | 13,3 | 16,2 | 18,2 | 21,1 | 23,7 | 26,9 | 29,1 | 31,3 | 34,0 | 36,1 |
| 15 | 5,2 | 6,0 | 7,3 | 8,5 | 10,3 | 11,7 | 14,3 | 17,3 | 19,3 | 22,3 | 25,0 | 28,3 | 30,6 | 32,8 | 35,6 | 37,7 |
| 16 | 5,8 | 6,6 | 8,0 | 9,3 | 11,2 | 12,6 | 15,3 | 18,4 | 20,5 | 23,5 | 26,3 | 29,6 | 32,0 | 34,3 | 37,1 | 39,3 |
| 17 | 6,4 | 7,3 | 8,7 | 10,1 | 12,0 | 13,5 | 16,3 | 19,5 | 21,6 | 24,8 | 27,6 | 31,0 | 33,4 | 35,7 | 38,6 | 40,8 |
| 18 | 7,0 | 7,9 | 9,4 | 10,9 | 12,9 | 14,4 | 17,3 | 20,6 | 22,8 | 26,0 | 28,9 | 32,3 | 34,8 | 37,2 | 40,1 | 42,3 |
| 19 | 7,6 | 8,6 | 10,1 | 11,7 | 13,7 | 15,4 | 18,3 | 21,7 | 23,9 | 27,2 | 30,1 | 33,7 | 36,2 | 38,6 | 41,6 | 43,8 |
| 20 | 8,3 | 9,2 | 10,9 | 12,4 | 14,6 | 16,3 | 19,3 | 22,8 | 25,0 | 28,4 | 31,4 | 35,0 | 37,6 | 40,0 | 43,0 | 45,3 |
| 21 | 8,9 | 9,9 | 11,6 | 13,2 | 15,4 | 17,2 | 20,3 | 23,9 | 26,2 | 29,6 | 32,7 | 36,3 | 38,9 | 41,4 | 44,5 | 46,8 |
| 22 | 9,5 | 10,6 | 12,3 | 14,0 | 16,3 | 18,1 | 21,3 | 24,9 | 27,3 | 30,8 | 33,9 | 37,7 | 40,3 | 42,8 | 45,9 | 48,3 |
| 23 | 10,2 | 11,3 | 13,1 | 14,8 | 17,2 | 19,0 | 22,3 | 26,0 | 28,4 | 32,0 | 35,2 | 39,0 | 41,6 | 44,2 | 47,3 | 49,7 |
| 24 | 10,9 | 12,0 | 13,8 | 15,7 | 18,1 | 19,9 | 23,3 | 27,1 | 29,6 | 33,2 | 36,4 | 40,3 | 43,0 | 45,6 | 48,7 | 51,2 |
| 25 | 11,5 | 12,7 | 14,6 | 16,5 | 18,9 | 20,9 | 24,3 | 28,2 | 30,7 | 34,4 | 37,7 | 41,6 | 44,3 | 46,9 | 50,1 | 52,6 |
| 26 | 12,2 | 13,4 | 15,4 | 17,3 | 19,8 | 21,8 | 25,3 | 29,2 | 31,8 | 35,6 | 38,9 | 42,9 | 45,6 | 48,3 | 51,6 | 54,1 |
| 27 | 12,9 | 14,1 | 16,2 | 18,1 | 20,7 | 22,7 | 26,3 | 30,3 | 32,9 | 36,7 | 40,1 | 44,1 | 47,0 | 49,6 | 52,9 | 55,5 |
| 28 | 13,6 | 14,8 | 16,9 | 18,9 | 21,6 | 23,6 | 27,3 | 31,4 | 34,0 | 37,9 | 41,3 | 45,4 | 48,3 | 51,0 | 54,4 | 56,9 |

| Anzahl der Freiheits-grade $m$ | Wahrscheinlichkeit $p = \alpha$ | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0,99 | 0,98 | 0,95 | 0,90 | 0,80 | 0,70 | 0,50 |
| 1 | 0,00016 | 0,0006 | 0,0039 | 0,016 | 0,064 | 0,148 | 0,455 |
| 2 | 0,020 | 0,040 | 0,103 | 0,211 | 0,446 | 0,713 | 1,386 |
| 3 | 0,115 | 0,185 | 0,352 | 0,584 | 1,005 | 1,424 | 2,366 |
| 4 | 0,30 | 0,43 | 0,71 | 1,06 | 1,65 | 2,19 | 3,36 |
| 5 | 0,55 | 0,75 | 1,14 | 1,61 | $\boxed{2,34}$ | 3,00 | 4,35 |
| 6 | 0,87 | 1,13 | 1,63 | 2,20 | 3,07 | 3,83 | 5,35 |

# Pearson's $\chi^2$ Test of Independence

- Pearson's $\chi^2$ test can also be used to test if two random variables $X$ and $Y$ are independent
- Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be the two samples
- Divide outcomes into $r$ (for $X$) and $c$ (for $Y$) disjoint intervals
- Populate $r$-by-$c$ table O with frequencies, so that $O_{lk}$ tells how many $(X_i, Y_i)$ pairs have values $l$-th and $k$-interval respectively
- Assuming independence ($H_0$) the expected value of $O_{lk}$ is

$$E_{lk} = \frac{\sum_{i=1}^{c} O_{li} \sum_{j=1}^{r} O_{jk}}{\sum_{j=1}^{r} \sum_{i=1}^{c} O_{ij}}$$

# Pearson's $\chi^2$ Test of Independence (cont'd)

- The value of the test statistic is

$$\chi^2 = \sum_{i=1}^{c} \sum_{j=1}^{r} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

- There are ($r$-1)($c$-1) degrees of freedom

# Summary of II.3

- **Statistical inference** based on a **sample** from a **population**
- **Empirical distribution function** and **histograms** as non-parametric estimation methods
- **Method of moments** and **maximum likelihood** as parametric estimation methods
- **Confidence intervals**
- **Wald test** and **Pearson's $\chi^2$ test** for **hypothesis testing**

# Normal Distribution Table

The Normal Distribution Functions $\Phi(z) = \int_{-\infty}^{z} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt$


| z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | .50000 | .50399 | .50798 | .51197 | .51595 | .51994 | .52392 | .52790 | .53188 | .53586 |
| 0.1 | .53983 | .54380 | .54776 | .55172 | .55567 | .55962 | .56356 | .56749 | .57142 | .57535 |
| 0.2 | .57926 | .58317 | .58706 | .59095 | .59483 | .59871 | .60257 | .60642 | .61026 | .61409 |
| 0.3 | .61791 | .62172 | .62552 | .62930 | .63307 | .63683 | .64058 | .64431 | .64803 | .65173 |
| 0.4 | .65542 | .65910 | .66276 | .66640 | .67003 | .67364 | .67724 | .68082 | .68439 | .68793 |
| 0.5 | .69146 | .69497 | .69847 | .70194 | .70540 | .70884 | .71226 | .71566 | .71904 | .72240 |
| 0.6 | .72575 | .72907 | .73237 | .73565 | .73891 | .74215 | .74537 | .74857 | .75175 | .75490 |
| 0.7 | .75804 | .76115 | .76424 | .76730 | .77035 | .77337 | .77637 | .77935 | .78230 | .78524 |
| 0.8 | .78814 | .79103 | .79389 | .79673 | .79955 | .80234 | .80511 | .80785 | .81057 | .81327 |
| 0.9 | .81594 | .81859 | .82121 | .82381 | .82639 | .82894 | .83147 | .83398 | .83646 | .83891 |
| 1.0 | .84134 | .84375 | .84614 | .84849 | .85083 | .85314 | .85543 | .85769 | .85993 | .86214 |
| 1.1 | .86433 | .86650 | .86864 | .87076 | .87286 | .87493 | .87698 | .87900 | .88100 | .88298 |
| 1.2 | .88493 | .88686 | .88877 | .89065 | .89251 | .89435 | .89617 | .89796 | .89973 | .90147 |
| 1.3 | .90320 | .90490 | .90658 | .90824 | .90988 | .91149 | .91308 | .91466 | .91621 | .91774 |
| 1.4 | .91924 | .92073 | .92220 | .92364 | .92507 | .92647 | .92785 | .92922 | .93056 | .93189 |
| 1.5 | .93319 | .93448 | .93574 | .93699 | .93822 | .93943 | .94062 | .94179 | .94295 | .94408 |
| 1.6 | .94520 | .94630 | .94738 | .94845 | .94950 | .95053 | .95154 | .95254 | .95352 | .95449 |
| 1.7 | .95543 | .95637 | .95728 | .95818 | .95907 | .95994 | .96080 | .96164 | .96246 | .96327 |
| 1.8 | .96407 | .96485 | .96562 | .96638 | .96712 | .96784 | .96856 | .96926 | .96995 | .97062 |
| 1.9 | .97128 | .97193 | .97257 | .97320 | .97381 | .97441 | .97500 | .97558 | .97615 | .97670 |
| 2.0 | .97725 | .97778 | .97831 | .97882 | .97932 | .97982 | .98030 | .98077 | .98124 | .98169 |
| 2.1 | .98214 | .98257 | .98300 | .98341 | .98382 | .98422 | .98461 | .98500 | .98537 | .98574 |
| 2.2 | .98610 | .98645 | .98679 | .98713 | .98745 | .98778 | .98809 | .98840 | .98870 | .98899 |
| 2.3 | .98928 | .98956 | .98983 | .99010 | .99036 | .99061 | .99086 | .99111 | .99134 | .99158 |
| 2.4 | .99180 | .99202 | .99224 | .99245 | .99266 | .99286 | .99305 | .99324 | .99343 | .99361 |
| 2.5 | .99379 | .99396 | .99413 | .99430 | .99446 | .99461 | .99477 | .99492 | .99506 | .99520 |
| 2.6 | .99534 | .99547 | .99560 | .99573 | .99585 | .99598 | .99609 | .99621 | .99632 | .99643 |
| 2.7 | .99653 | .99664 | .99674 | .99683 | .99693 | .99702 | .99711 | .99720 | .99728 | .99736 |
| 2.8 | .99744 | .99752 | .99760 | .99767 | .99774 | .99781 | .99788 | .99795 | .99801 | .99807 |
| 2.9 | .99813 | .99819 | .99825 | .99831 | .99836 | .99841 | .99846 | .99851 | .99856 | .99861 |
| 3.0 | .99865 | .99869 | .99874 | .99878 | .99882 | .99886 | .99889 | .99893 | .99896 | .99900 |
| 3.1 | .99903 | .99906 | .99910 | .99913 | .99916 | .99918 | .99921 | .99924 | .99926 | .99929 |
| 3.2 | .99931 | .99934 | .99936 | .99938 | .99940 | .99942 | .99944 | .99946 | .99948 | .99950 |
| 3.3 | .99952 | .99953 | .99955 | .99957 | .99958 | .99960 | .99961 | .99962 | .99964 | .99965 |
| 3.4 | .99966 | .99968 | .99969 | .99970 | .99971 | .99972 | .99973 | .99974 | .99975 | .99976 |
| 3.5 | .99977 | .99978 | .99978 | .99979 | .99980 | .99981 | .99981 | .99982 | .99983 | .99983 |
| 3.6 | .99984 | .99985 | .99985 | .99986 | .99986 | .99987 | .99987 | .99988 | .99988 | .99989 |
| 3.7 | .99989 | .99990 | .99990 | .99990 | .99991 | .99991 | .99992 | .99992 | .99992 | .99992 |
| 3.8 | .99993 | .99993 | .99993 | .99994 | .99994 | .99994 | .99994 | .99995 | .99995 | .99995 |

# $\chi^2$ Distribution Table

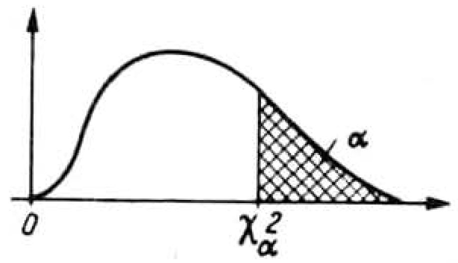
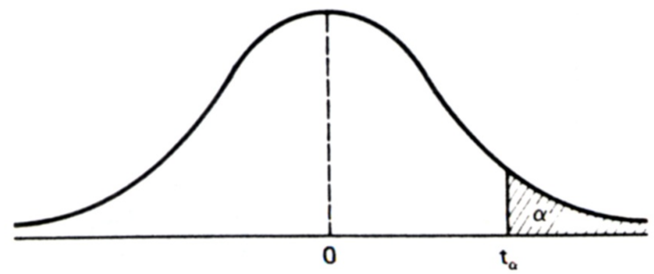



Abb. 1.4

| Anzahl der Freiheitsgrade $m$ | Wahrscheinlichkeit $p = \alpha$ | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0,99 | 0,98 | 0,95 | 0,90 | 0,80 | 0,70 | 0,50 | 0,30 | 0,20 | 0,10 | 0,05 | 0,02 | 0,01 | 0,005 | 0,002 | 0,001 |
| 1 | 0,00016 | 0,0006 | 0,0039 | 0,016 | 0,064 | 0,148 | 0,455 | 1,07 | 1,64 | 2,7 | 3,8 | 5,4 | 6,6 | 7,9 | 9,5 | 10,83 |
| 2 | 0,020 | 0,040 | 0,103 | 0,211 | 0,446 | 0,713 | 1,386 | 2,41 | 3,22 | 4,6 | 6,0 | 7,8 | 9,2 | 10,6 | 12,4 | 13,8 |
| 3 | 0,115 | 0,185 | 0,352 | 0,584 | 1,005 | 1,424 | 2,366 | 3,67 | 4,64 | 6,3 | 7,8 | 9,8 | 11,3 | 12,8 | 14,8 | 16,3 |
| 4 | 0,30 | 0,43 | 0,71 | 1,06 | 1,65 | 2,19 | 3,36 | 4,9 | 6,0 | 7,8 | 9,5 | 11,7 | 13,3 | 14,9 | 16,9 | 18,5 |
| 5 | 0,55 | 0,75 | 1,14 | 1,61 | 2,34 | 3,00 | 4,35 | 6,1 | 7,3 | 9,2 | 11,1 | 13,4 | 15,1 | 16,8 | 18,9 | 20,5 |
| 6 | 0,87 | 1,13 | 1,63 | 2,20 | 3,07 | 3,83 | 5,35 | 7,2 | 8,6 | 10,6 | 12,6 | 15,0 | 16,8 | 18,5 | 20,7 | 22,5 |
| 7 | 1,24 | 1,56 | 2,17 | 2,83 | 3,82 | 4,67 | 6,35 | 8,4 | 9,8 | 12,0 | 14,1 | 16,6 | 18,5 | 20,3 | 22,6 | 24,3 |
| 8 | 1,65 | 2,03 | 2,73 | 3,49 | 4,59 | 5,53 | 7,34 | 9,5 | 11,0 | 13,4 | 15,5 | 18,2 | 20,1 | 22,0 | 24,3 | 26,1 |
| 9 | 2,09 | 2,53 | 3,32 | 4,17 | 5,38 | 6,39 | 8,34 | 10,7 | 12,2 | 14,7 | 16,9 | 19,7 | 21,7 | 23,6 | 26,1 | 27,9 |
| 10 | 2,56 | 3,06 | 3,94 | 4,86 | 6,18 | 7,27 | 9,34 | 11,8 | 13,4 | 16,0 | 18,3 | 21,2 | 23,2 | 25,2 | 27,7 | 29,6 |
| 11 | 3,1 | 3,6 | 4,6 | 5,6 | 7,0 | 8,1 | 10,3 | 12,9 | 14,6 | 17,3 | 19,7 | 22,6 | 24,7 | 26,8 | 29,4 | 31,3 |
| 12 | 3,6 | 4,2 | 5,2 | 6,3 | 7,8 | 9,0 | 11,3 | 14,0 | 15,8 | 18,5 | 21,0 | 24,1 | 26,2 | 28,3 | 30,9 | 32,9 |
| 13 | 4,1 | 4,8 | 5,9 | 7,0 | 8,6 | 9,9 | 12,3 | 15,1 | 17,0 | 19,8 | 22,4 | 25,5 | 27,7 | 29,8 | 32,5 | 34,5 |
| 14 | 4,7 | 5,4 | 6,6 | 7,8 | 9,5 | 10,8 | 13,3 | 16,2 | 18,2 | 21,1 | 23,7 | 26,9 | 29,1 | 31,3 | 34,0 | 36,1 |
| 15 | 5,2 | 6,0 | 7,3 | 8,5 | 10,3 | 11,7 | 14,3 | 17,3 | 19,3 | 22,3 | 25,0 | 28,3 | 30,6 | 32,8 | 35,6 | 37,7 |
| 16 | 5,8 | 6,6 | 8,0 | 9,3 | 11,2 | 12,6 | 15,3 | 18,4 | 20,5 | 23,5 | 26,3 | 29,6 | 32,0 | 34,3 | 37,1 | 39,3 |
| 17 | 6,4 | 7,3 | 8,7 | 10,1 | 12,0 | 13,5 | 16,3 | 19,5 | 21,6 | 24,8 | 27,6 | 31,0 | 33,4 | 35,7 | 38,6 | 40,8 |
| 18 | 7,0 | 7,9 | 9,4 | 10,9 | 12,9 | 14,4 | 17,3 | 20,6 | 22,8 | 26,0 | 28,9 | 32,3 | 34,8 | 37,2 | 40,1 | 42,3 |
| 19 | 7,6 | 8,6 | 10,1 | 11,7 | 13,7 | 15,4 | 18,3 | 21,7 | 23,9 | 27,2 | 30,1 | 33,7 | 36,2 | 38,6 | 41,6 | 43,8 |
| 20 | 8,3 | 9,2 | 10,9 | 12,4 | 14,6 | 16,3 | 19,3 | 22,8 | 25,0 | 28,4 | 31,4 | 35,0 | 37,6 | 40,0 | 43,0 | 45,3 |
| 21 | 8,9 | 9,9 | 11,6 | 13,2 | 15,4 | 17,2 | 20,3 | 23,9 | 26,2 | 29,6 | 32,7 | 36,3 | 38,9 | 41,4 | 44,5 | 46,8 |
| 22 | 9,5 | 10,6 | 12,3 | 14,0 | 16,3 | 18,1 | 21,3 | 24,9 | 27,3 | 30,8 | 33,9 | 37,7 | 40,3 | 42,8 | 45,9 | 48,3 |
| 23 | 10,2 | 11,3 | 13,1 | 14,8 | 17,2 | 19,0 | 22,3 | 26,0 | 28,4 | 32,0 | 35,2 | 39,0 | 41,6 | 44,2 | 47,3 | 49,7 |
| 24 | 10,9 | 12,0 | 13,8 | 15,7 | 18,1 | 19,9 | 23,3 | 27,1 | 29,6 | 33,2 | 36,4 | 40,3 | 43,0 | 45,6 | 48,7 | 51,2 |
| 25 | 11,5 | 12,7 | 14,6 | 16,5 | 18,9 | 20,9 | 24,3 | 28,2 | 30,7 | 34,4 | 37,7 | 41,6 | 44,3 | 46,9 | 50,1 | 52,6 |
| 26 | 12,2 | 13,4 | 15,4 | 17,3 | 19,8 | 21,8 | 25,3 | 29,2 | 31,8 | 35,6 | 38,9 | 42,9 | 45,6 | 48,3 | 51,6 | 54,1 |
| 27 | 12,9 | 14,1 | 16,2 | 18,1 | 20,7 | 22,7 | 26,3 | 30,3 | 32,9 | 36,7 | 40,1 | 44,1 | 47,0 | 49,6 | 52,9 | 55,5 |
| 28 | 13,6 | 14,8 | 16,9 | 18,9 | 21,6 | 23,6 | 27,3 | 31,4 | 34,0 | 37,9 | 41,3 | 45,4 | 48,3 | 51,0 | 54,4 | 56,9 |

# Student's *t* Distribution Table




| $n$ \ $\alpha$ | 0.10 | 0.05 | 0.025 | 0.01 | 0.005 |
|---|---|---|---|---|---|
| 1 | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 |
| 2 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 |
| 3 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 |
| 4 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 |
| 5 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 |
| 6 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 |
| 8 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 |
| 9 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 |
| 10 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 |
| 11 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 |
| 12 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 |
| 13 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 |
| 14 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 |
| 15 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 |
| 16 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 |
| 17 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 |
| 18 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 |
| 19 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 |
| 20 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 |
| 21 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 |
| 22 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 |
| 23 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 |
| 24 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 |
| 25 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 |
| 26 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 |
| 27 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 |
| 28 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 |
| 29 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 |
| 30 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 |
| 40 | 1.303 | 1.684 | 2.021 | 2.423 | 2.704 |
| 60 | 1.296 | 1.671 | 2.000 | 2.390 | 2.660 |
| 120 | 1.289 | 1.658 | 1.980 | 2.358 | 2.617 |
| $\infty$ | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 |