1 R Tutorial

1.1 Loading Data

```r
# loading csv files
data <- read.table("whatever.csv", sep="", header=T)

# csv files can be stored with (almost) any kind of file ending, e.g.:
data <- read.table("whatever.dat", sep="", header=T)
data <- read.table("whatever.txt", sep="", header=T)
```

2 Probability And Statistics

2.1 Probability Models for Measurement Data

2.1.1 Random Variables

<table>
<tr><th colspan="2">Random Variables</th></tr>
<tr><td>Definition</td><td>$X : \Omega \longrightarrow W_{\mathrm{x}}$</td></tr>
<tr><td>Example</td><td>A Coin is thrown three times, head and tails is observed:
$\Omega = \{hhh, hht, htt, hth, ttt, tth, thh, tht\}$
Total number of heads $W_{\mathrm{x}} = \{0, 1, 2, 3\}$
Total number of tails $W_{\mathrm{x}} = \{0, 1, 2, 3\}$
Number of heads minus tails $W_{\mathrm{x}} = \{-3, -1, 1, 3\}$</td></tr>
<tr><th colspan="2">Probability Mass Function</th></tr>
<tr><td>Definition</td><td>The probability distribution of a discrete random variable:
$P(X = x)$</td></tr>
<tr><td>Example</td><td><table><tr><td>x</td><td>0</td><td>1</td><td>2</td><td>3</td></tr><tr><td>$P(X = x)$</td><td>$\frac{1}{8}$</td><td>$\frac{3}{8}$</td><td>$\frac{3}{8}$</td><td>$\frac{1}{8}$</td></tr></table></td></tr>
</table>

2.1.2 Probability Distributions

<table>
<tr><th colspan="2">Cumulative Density Function (cdf)</th></tr>
<tr><td>Definition</td><td>$F(x) = P(X \leqslant x)$</td></tr>
<tr><td>Properties</td><td>$P(a < X \leqslant b) = F(b) - F(a)$
$0 \leqslant F(x) \leqslant 1$
$P(X = a) = F(a) - F(a) = 0$</td></tr>
</table>

Probability Density Function (pdf)	
Definition	$f(x) = \frac{\mathrm{d}F(x)}{\mathrm{d}x}$
Properties	$f(x) \geqslant 0$ $P(a < X \leqslant b) = F(b) - F(a) = \int_a^b f(x)\mathrm{d}x$ $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$

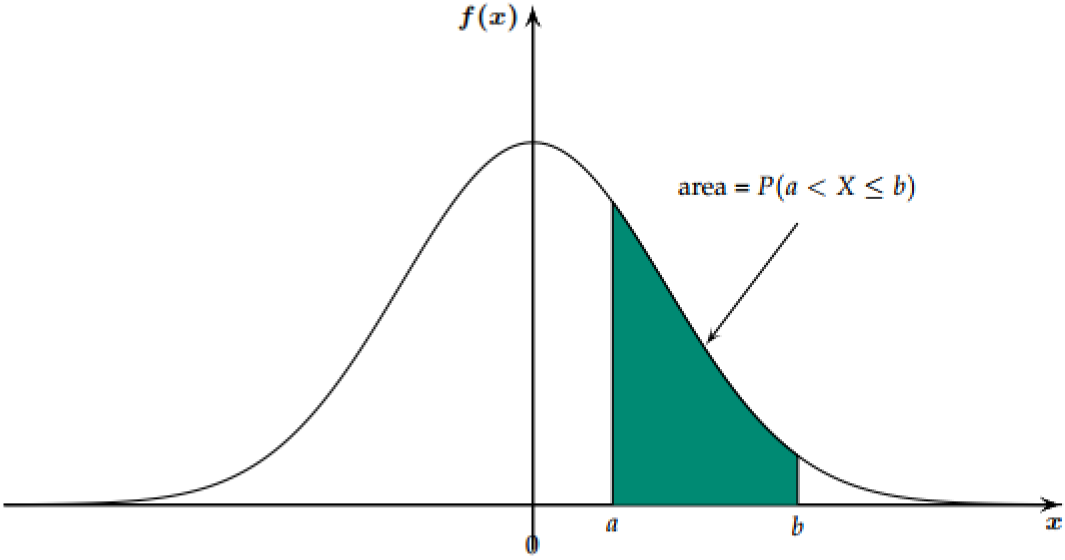


Figure 1: Probability density of a random variable and the probability of measuring a value from (a,b]

2.1.3 Summary Statistics of Continuous Distributions

Expected Value, Variance and Quantile	
Expected value	Discrete: $\mathrm{E}(X) = \sum_i x_i P(X = x_i)$ Continuous: $\mathrm{E}(X) = \mu_X = \int_{-\infty}^{\infty} x \cdot f(x)\mathrm{d}x$
Variance	$\mathrm{Var}(X) = \sigma_x^2 = \mathrm{E}((X - \mathrm{E}(X))^2) = \int_{-\infty}^{\infty} (x - \mathrm{E}(X))^2 \cdot f(x)\mathrm{d}x$
Quantile	$P(X \leqslant q(\alpha)) = \alpha$ $F(q(\alpha)) = \alpha \Leftrightarrow q(\alpha) = F^{-1}(\alpha)$ Note: When you're asked for the 50%-quantile, that means $\alpha = 50\%$, and you must find $q(0.5)$
Example Body Length	If α=0.75 and the corresponding quantile is $q(\alpha)$=182.5cm then 75% of the persons is shorter or equal 182.5cm.

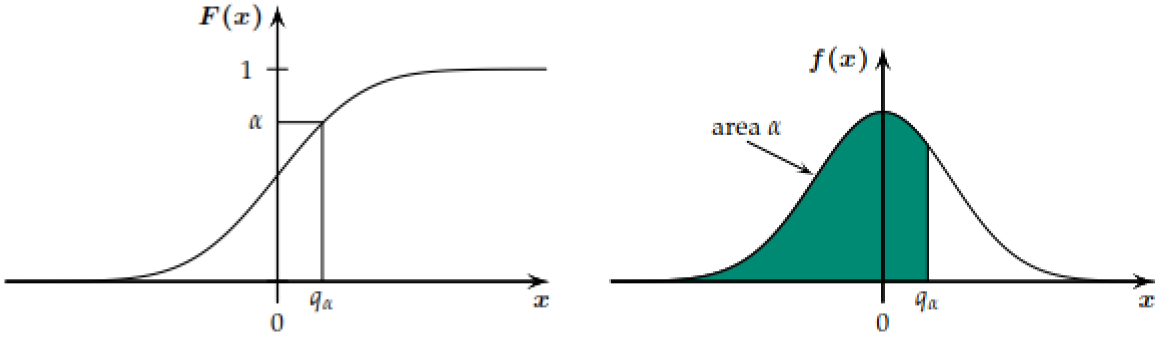


Figure 2: Quantiles

2.1.4 Important Distributions

2.1.4.1 Uniform Distribution

Theory

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leqslant x \leqslant b \\ 0 & \text{otherwise} \end{cases}$$

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leqslant x \leqslant b \\ 1 & x > b \end{cases}$$

$$\mathrm{E}(x) = \frac{a+b}{2}$$

$$\mathrm{Var}(x) = \frac{(b-a)^2}{12}$$

$$\sigma_{\mathrm{x}} = \frac{b-a}{\sqrt{12}}$$

Code Example

```r
# value of the probability density function
    Uniform([1, 10]) at the position x = 5
dunif(x=5, min=1, max=10)
[1] 0.1111111

# P(X <= 5)
punif(q=5, min=1, max=10)
[1] 0.4444444

# P(1.2 < X <= 4.8)
punif(4.8, 1, 10) - punif(1.2, 1, 10)
[1] 0.4

# 5 uniformly distributed random values in
    Uniform([1, 10])
runif(5,min=1,max=10)
[1] 1.061933 6.484813 5.928334 8.459887 8.852405

# TODO: ADD MORE HERE
```

2.1.4.2 Exponential Distribution

Theory

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x} & x \geqslant 0 \\ 0 & \text{otherwise} \end{cases}$$

$$F(x) = \begin{cases} 1 - \lambda \cdot e^{-\lambda \cdot x} & x \geqslant 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\mathrm{E}(x) = \frac{1}{\lambda}$$

$$\mathrm{Var}(x) = \frac{1}{\lambda^2}$$

$$\sigma_{\mathrm{x}} = \frac{1}{\lambda}$$

Code Example

```r
# P(0 <= X <= 4) of X ~ Exp(3)
pexp(4, rate=3)
[1] 0.9999939

# TODO: ADD MORE HERE
```

2.1.4.3 Normal Distribution

Theory	Code Example

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{(-\frac{(x-\mu)^2}{2\sigma^2})}$$ $$F(x) = \int_{-\infty}^{x} f(x)\mathrm{d}y$$ $$\mathrm{E}(x) = \mu$$ $$\mathrm{Var}(x) = \sigma^2$$ $$\sigma_{\mathrm{x}} = \sigma$$	`# X~N(u, sigma^2) --> X~N(100,15^2)`
	`# In R we compute P(X>130) as 1 - P(X<=130)`
	`1-pnorm(130, mean=100, sd=15)`
	`[1] 0.02275013`
	
	`#P(85<=X<=115)`
	`pnorm(115, mean=100, sd=15)-pnorm(85, mean=100, sd=15)`
	`[1] 0.6826895`
	
	`# TODO: ADD MORE HERE`

2.1.4.4 Linear Transformation of Random Variables

Properties of Linear Transformation of a Random Variable	
Definition	For $Y = a + bX$ the following apply (i) $E(Y) = a + bE(X)$ (ii) $Var(Y) = b^2 Var(X)$, $\sigma_Y = \lvert b\rvert\sigma_X$ (iii) $\alpha - Quantile\,of\,Y = q_Y(\alpha) = a + bq_X(\alpha)$ (iv) $f_Y(y) = \frac{1}{b} f_X(\frac{y-a}{b})$
Summary Statistics of S_n and $\bar{X}_n$	
Summary Statistics of Sample Total S_n	$E(S_n) = E(X_1 + X_2 + ... + X_n) = \sum_{i=1}^{n} E(X_i) = n\mu$ $Var(S_n) = \sum_{i=1}^{n} Var(X_i) = nVar(X_i)$ $\sigma(S_n) = \sqrt{n}\sigma_X$
Summary Statistics of Sample Mean $\bar{X}_n$ *Standard Error*	$E(\bar{X}_n) = E(\frac{X_1+X_2+...+X_n}{n}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} nE(X_i) = \mu$ $Var(\bar{X}_n) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2} n\sigma_X^2 = \frac{\sigma_X^2}{n}$ $\sigma(\bar{X}_n) = \frac{\sigma_X}{\sqrt{n}}$

2.1.4.5 Distributions of S_n and $\bar{X}_n$

Theory	Code Example

1. For $X_i \in \{0, 1\}$, we have $$S_n \sim \mathrm{Bin}(n, \pi) \text{ with } \pi = P(X_i = 1)$$ 2. For $X_i \sim \mathrm{Pois}(\lambda)$, we have $$S_n \sim \mathrm{Pois}(n\lambda)$$ 3. For $X_i \sim N(\mu, \sigma^2)$ $$S_n \sim N(n\mu, n\sigma^2) \text{ and } \bar{X}_n \sim N(\mu, \frac{\sigma_X^2}{n})$$	`What is the probability that among 10000 tosses of a fair coin, heads would appear in maximum 5100 cases?`
	`#Approximated: X~N(5000,2500)`
	`pnorm(5100,5000,sqrt(2500))`
	`[1] 0.9772499`
	
	`#"True Result": X~Bin(10000,0.5)`
	`pbinom(5100,10000,0.5)`
	`[1] 0.9777871`

2.2 Statistics for Measurement Data

bgg

2.2.1 Assess the Normal Distribution Assumption

2.2.1.1 Q-Q Plot

Theory

1. For

$$\alpha_k = \frac{k-0.5}{n} \text{ with } k = 1, ..., n$$

calculate the corresponding theoretical quantiles of the model distribution

$$q(\alpha_k) = F^{-1}(\alpha_k)$$

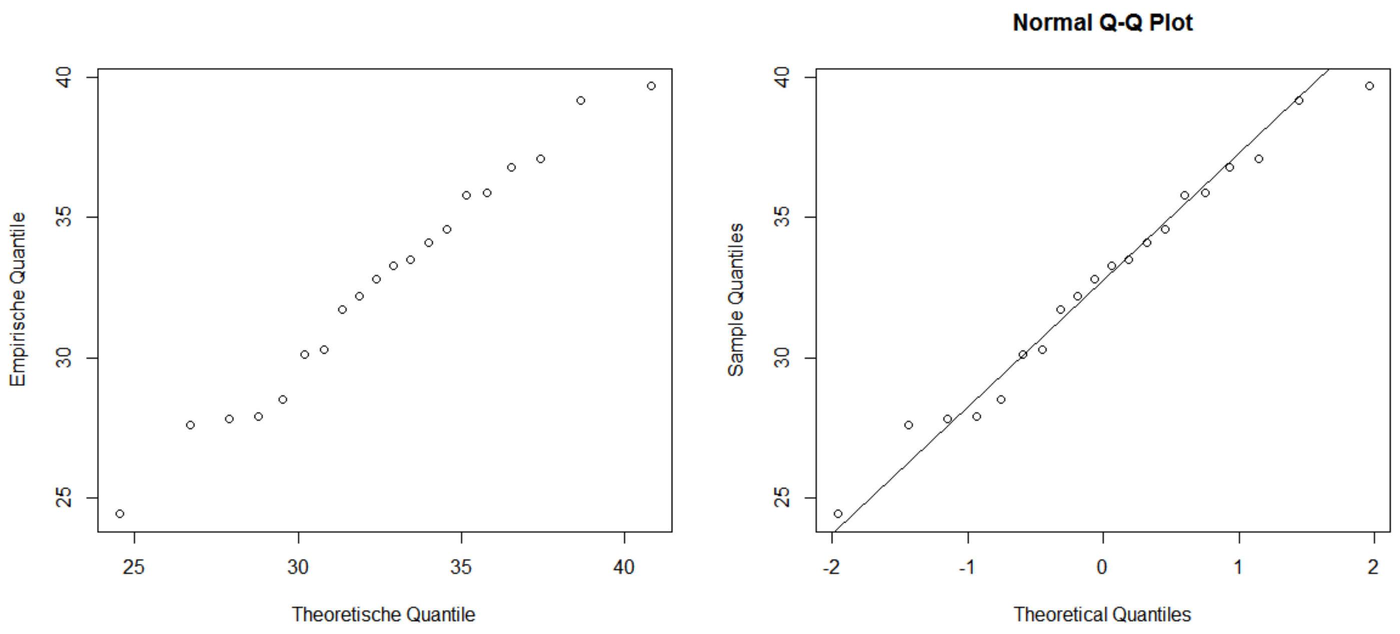
2. Determine the empirical α_k-quantiles,

$$x_{(1)} < x_{(2)} < ... < x_{(n)}$$

3. Plot the empirical quantiles x_k on the y-axis against the theoretical quantiles $q(\alpha_k)$ on the x-axis.

Code Example

```r
x <- c(24.4, 27.6, 27.8, 27.9, 28.5, 30.1, 30.3,
    31.7, 32.2, 32.8, 33.3, 33.5, 34.1, 34.6,
    35.8, 35.9, 36.8, 37.1, 39.2, 39.7)

alpha_k <- (seq(1, length(x), by=1)-0.5)/length(
    x)

quantile_th <- qnorm(alpha_k, mean=mean(x), sd=
    sd(x))

quantile_emp <- sort(x)
#image qqplot
qqplot(quantile_th, quantile_emp, xlab="
    Theoretische Quantile", ylab = "Empirische
    Quantile")
#image qqnorm;qqline
qqnorm(x);qqline(x)
```



(a) qqplot()

(b) qqnorm();qqline()

k	$x_{(k)}$	$\alpha_k = (k-0.5)/n$	q_{α_k} for $\mathcal{N}(32.7, 4.15^2)$	$\Phi^{-1}(\alpha_k)$
1	24.4	0.0250	24.5	-1.96
2	27.6	0.075	26.7	-1.44
3	27.8	0.125	27.9	-1.15
4	27.9	0.175	28.8	-0.935
5	28.5	0.225	29.5	-0.755
6	30.1	0.275	30.2	-0.600
7	30.3	0.325	30.8	-0.453
8	31.7	0.375	31.3	-0.319
9	32.2	0.425	31.9	-0.189
10	32.8	0.475	32.4	-0.0627
11	33.3	0.525	32.9	0.0627
12	33.5	0.575	33.4	0.189
13	34.1	0.625	34.0	0.319
14	34.6	0.675	34.5	0.454
15	35.8	0.725	35.1	0.598
16	35.9	0.775	36.0	0.755
17	36.8	0.825	36.5	0.935
18	37.1	0.875	37.4	1.15
19	39.2	0.925	38.6	1.44
20	39.7	0.975	40.8	1.96

```r
#x(k) are the measured values N(u,sigma^2)
x <- c(24.4, 27.6, 27.8, 27.9, 28.5, 30.1, 30.3,
    31.7, 32.2, 32.8, 33.3, 33.5, 34.1, 34.6, 35.8,
     35.9, 36.8, 37.1, 39.2, 39.7)
mean(x)
[1] 32.665
sd(x)
[1] 4.149734
#N(32.7,4.15)
#a_k = (k-0.5)/n = qnorm(q_ak, 32.7, 4.15)
pnorm(24.5, 32.7, 4.15)
[1] 0.02408285
pnorm(32.4, 32.7, 4.15)
[1] 0.4711859
pnorm(35.8, 32.7, 4.15)
[1] 0.7724646
pnorm(40.8, 32.7, 4.15)
[1] 0.9745195
#q_ak for N(32.7,4.15) = qnorm(a_k, 32.7, 4.15)
qnorm(0.025, 32.7, 4.15)
[1] 24.56615
qnorm(0.475, 32.7, 4.15)
[1] 32.43977
qnorm(0.725, 32.7, 4.15)
[1] 35.1807
qnorm(0.975, 32.7, 4.15)
[1] 40.83385
#phi^{-1}(a_k)
qnorm(0.025)
[1] -1.959964
qnorm(0.475)
[1] -0.06270678
qnorm(0.725)
[1] 0.5977601
qnorm(0.975)
[1] 1.959964
```

2.2.2 Parameter Esitmation for Continuous Probability Distributions

Method of Moments (not unbiased)

1. We consider our data measurements $x_1, x_2, ..., x_n$ as realization of random variables $X_1, X_2, ..., X_n$ originating from the same known distribution.

2. We calculate the expected value $E(X)$ and solve the equation for the unknown parameter that we intend to estimate.

3. We replace the expected value with its counterpart, the empirical mean value and obtain an estimate of the unknown parameter. A method of moments estimate of the standard deviation is the empirical standard deviation.

$$\mu = E(X) \Rightarrow \hat{\mu} = \bar{x}_n = \frac{x_1+x_2+...+x_n}{n} = \frac{653.3}{20} = 32.7$$

$$\sigma^2 = E(X^2) - E(X)^2 = E(X^2) - \mu^2$$

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

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

$$\hat{\sigma}^2 = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2} = \sqrt{\frac{1}{20}\sum_{i=1}^{20}(x_i - 32.7)^2} = 4.04$$

Method of Maximum Likelihood

We have n observations that are i.i.d.

$$X_1 = x_1, X_2 = x_2, ..., X_n = x_n$$

For a discrete probability distribution: probability that these n observations (events) actually have occurred can be expressed as follows

$$P[(X_1 = x_1) \cap (X_2 = x_2) \cap ... \cap (X_n = x_n)] =$$
$$P[X_1 = x_1] \cdot P[X_2 = x_2] \cdot ... \cdot P[X_n = x_n] = \prod_{i=1}^{n} P[X_i = x_i]$$

Probability that the n independent random variables $x_1, x_2, ..., x_n$ are observed, depends on parameter θ, which we wish to estimate. Therefore the Likelihood function is given by $L(\theta)$ where $P[X_i = x_i\|\theta]$ denotes probability mass function that value x_i has been observed, given the parameter value θ.	$L(\theta) = P[X_1 = x_1\|\theta] \cdot P[X_2 = x_2\|\theta] \cdot ... \cdot P[X_n = x_n\|\theta] = \prod_{i=1}^{n} P[X_i = x_i\|\theta]$
Idea of Maximum Likelihood : estimate the parameter θ in such a way that the likelihood is maximized, that is, that it makes the observed data most likely or most probable.	
Continuous probability distributions : with probability density function $f(x;\theta)$. Probability, that each observation x_i falls into its corresponding interval $[x_i, x_i + dx_i]$:	$\prod_{i=1}^{n} f(x_i;\theta)dx_i$
Infinitesimal intervals dx_i do not depend on the parameter value θ : we omit them in the likelihood function	$\prod_{i=1}^{n} f(x_i;\theta)$
If assumed probability density function $f(x_i;\theta)$ and parameter value of θ are correct, we expect a high probability for the actually observed data to occur : maximization of $L(\theta)$	

Example: Maximum Likelihood for Exponential Distribution	
Let $X_1, X_2..., X_n$ i.i.d. $\sim$ Exp(λ), that is	$f(x_i;\lambda) = \lambda e^{-\lambda x_i}$
Likelihood function for a given data set $x_1, x_2, ..., x_n$ is given by	$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$
Log likelihood function is	$\log(L(\lambda)) = n\log(\lambda) - \lambda \sum_{i=1}^{n} x_i$
If we calculate the derivative of the log likelihood function with respect to λ and set it equal to 0, then we obtain	$\frac{d\log(L(\lambda))}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i \overset{!}{=} 0$
The maximum likelihood estimate $\hat{\lambda}$ thus corresponds to the solution of the previous equation	$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{x}}$

2.2.3 Statistical Tests and Confidence Interval for Normally Distributed Data

z-Test (σ_x known)	
1. Model:	$X_1, ..., X_n$ i.i.d. $\sim N(\mu, \sigma_X^2)$, σ_X known
2. Null hypothesis: Alternative:	H_0: $\mu = \mu_0$ H_A: $\mu \neq \mu_0$ (or $<$ or $>$)
3. Test statistic: Null distribution (assuming H_0 is true):	$Z = \frac{(\bar{X}_n - \mu_0)}{\sigma_{\bar{X}_n}} = \frac{(\bar{X}_n - \mu_0)}{\sigma_{X_n}/\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\sigma_{X_n}} = \frac{observed - expected}{standard\, error}$ $Z \sim N(0, 1)$
4. Significance level:	α
5. Rejection region for the test statistic: where	$K = (-\infty, z_{\frac{\alpha}{2}}] \cup [z_{1-\frac{\alpha}{2}}, \infty)$ with $H_A : \mu \neq \mu_0$, $K = (-\infty, z_\alpha]$ with $H_A : \mu < \mu_0$, $K = [z_{1-\alpha}, \infty)$ with $H_A : \mu > \mu_0$ $z_{\frac{\alpha}{2}} = \Phi^{-1}(\alpha/2)$

6. Test decision:	Check whether the observed value of the test statistic falls into the rejection region.

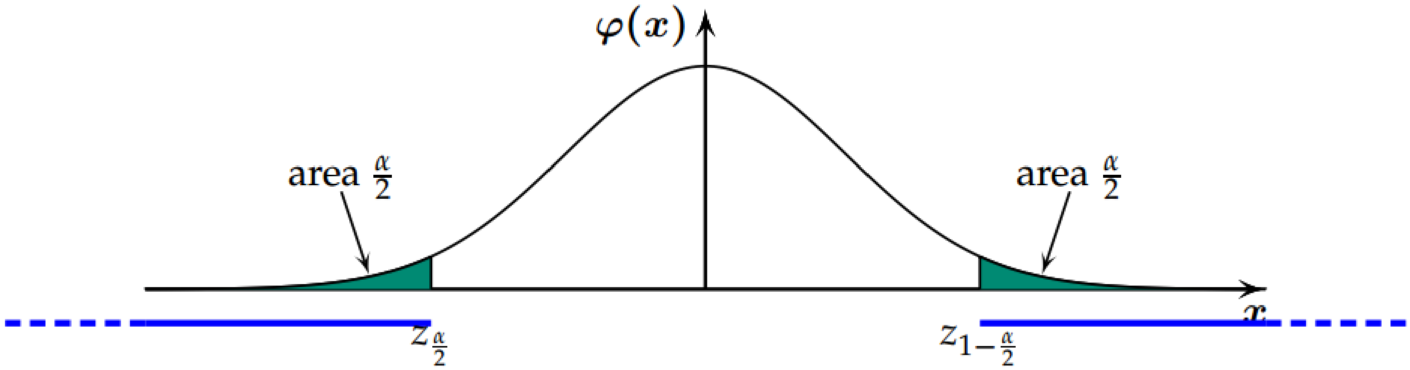


Figure 4: z-Test: Rejection Region

z-Test (σ_x known): Example	
Measurement of fusion heat:	The empirical mean value of $n = 13$ measurements is 80.02. From previous measurements the standard deviation is $\sigma_X = 0.01$. Is a fusion heat of exactly $80.00\frac{g}{cal}$ plausible?
1. Model:	$X_1, ..., X_n$ i.i.d. $\sim N(\mu, \sigma_X^2)$, $\sigma_X = 0.01$ known, $n = 13$
2. Null hypothesis: Alternative:	H_0: $\mu = \mu_0 = 80.00$ H_A: $\mu \neq \mu_0$
3. Test statistic: Null distribution (assuming H_0 is true):	$Z = \frac{\sqrt{n}\bar{X}_n - \mu_0)}{\sigma_{X_n}}$ $Z \sim N(0,1)$
4. Significance level:	$\alpha = 0.05$ (commonly used α-level)

5. Rejection region for the test statistic:
 	
Given $\alpha = 0.05$, R yields the following 2.5% quantile of the standard normal distribution.
 	
The following rejection region for the test statistic results
	$K = (-\infty, z_{\frac{\alpha}{2}}] \cup [z_{1-\frac{\alpha}{2}}, \infty)$ with $H_A: \mu \neq \mu_0$
	
	`qnorm(0.025)`
	`[1] -1.959964`
	
	$z_{\frac{\alpha}{2}} = \Phi^{-1}(\alpha/2) = \Phi^{-1}(0.025) = -1.96$
	$K = (-\infty, -1.96] \cup [1.96, \infty)$

6. Test decision: **Remarks:** Standardizing is in principle unnecessary because of technical aid of computer software.	Hence the value for the statistics is $z = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\sigma_{X_n}} = \frac{\sqrt{13}(80.02-80.00)}{0.01} = 7.211$ Therefore the observed value falls into the rejection region.
3. Test statistic: (not standardized)	The mean value of the measurements T: $\bar{X}_n$

Null distribution (assuming H_0 is true):	$T \sim N(\mu_0, \frac{\sigma_X^2}{n}) = N(80, \frac{0.01^2}{13})$

5. Rejection region for the test statistic: (not standardized)
 	
Given $\alpha = 0.05$, R yields the following 2.5% quantile of the standard normal distribution.
 	
In this way, we obtain the rejection region tor the test statistic:
	$K = (-\infty, c_u] \cup [c_o, \infty)$ with $H_A : \mu \neq \mu_0$
	
	`qnorm(0.025, 80.0, 0.01/sqrt(13))`
	`[1] 79.99456`
	`qnorm(0.975, 80.0, 0.01/sqrt(13))`
	`[1] 80.00544`
	
	$K = (-\infty, 79.99] \cup [80.01, \infty)$

***t*-Test (σ_x unknown)**	
1. Model:	$X_1, ..., X_n$ i.i.d. $\sim N(\mu, \sigma_X^2)$, σ_X is estimated by $\hat{\sigma}_X$
2. Null hypothesis: Alternative:	H_0: $\mu = \mu_0$ H_A: $\mu \neq \mu_0$ (or $<$ or $>$)
3. Test statistic: Null distribution (assuming H_0 is true):	$T = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\hat{\sigma}_X} = \frac{observed - expected}{estimated\ standard\ error}$ $T \sim t_{n-1}$
4. Significance level:	α
5. Rejection region for the test statistic:	$K = (-\infty, t_{n-1;\frac{\alpha}{2}}] \cup [t_{n-1;1-\frac{\alpha}{2}}, \infty)$ with $H_A : \mu \neq \mu_0$, $K = (-\infty, t_{n-1;\alpha}]$ with $H_A : \mu < \mu_0$, $K = [t_{n-1;1-\alpha}, \infty)$ with $H_A : \mu > \mu_0$
6. Test decision:	Check whether the observed value of the test statistic falls into the rejection region.
Example	
1. Model:	$X_1, ..., X_n$ i.i.d. $\sim N(\mu, \sigma_X^2)$, σ_X is estimated, $\hat{\sigma}_X = 0.024$
2. Null hypothesis: Alternative:	H_0: $\mu = \mu_0 = 80.00$ H_A: $\mu \neq \mu_0$
3. Test statistic: Null distribution (assuming H_0 is true):	$T = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\hat{\sigma}_X}$ $T \sim t_{n-1}$
4. Significance level:	$\alpha = 0.05$

5. Rejection region for the test statistic:
 	
We determine the value
 	
by means of R, where $\alpha = 0.05$ and $n = 13$.
 	
The rejection region of the test statistic thus is given by
	$K = (-\infty, t_{n-1;\frac{\alpha}{2}}] \cup [t_{n-1;1-\frac{\alpha}{2}}, \infty)$ with $H_A : \mu \neq \mu_0$,
	
	$t_{n-1;1-\frac{\alpha}{2}} = t_{12;0.975} = 2.179$
	
	`qt(0.975,12)`
	`[1] 2.178813`
	
	$K = (-\infty, -2.179] \cup [2.179, \infty)$

6. Test decision:	On the basis of $n = 13$ measurements, we find

$\bar{x} = 80.02$ and $\hat{\sigma}_X = 0.024$

Hence, the realized value of the test statistic is

$t = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\hat{\sigma}_X} = \frac{\sqrt{13}(80.02-80.00)}{0.024} = 3.00$

The observed value falls into the rejection region. Therefore, the null hypothesis is rejected at 5% level.

The t-test directly performed in R using the function t.test()

```
x <- c(79.98, 80.04, 80.02, 80.04, 80.03,
80.03, 80.04, 79.97, 80.05, 80.03,
80.02, 80.00, 80.02)

t.test(x, alternative = "two.sided",
mu = 80.00, conf.level = 0.95)

##
## One Sample t-test
##
## data:  x
## t = 3.1246, df = 12, p-value = 0.008779
## alternative hypothesis: true mean is not
    equal to 80
## 95 percent confidence interval:
## 80.00629 80.03525
## sample estimates:
## mean of x
## 80.02077
```

Remarks:
(i) The observed value of the test statistic is 3.12. Assuming the null hypothesis is true, then the test statistic follows a t-distribution with $df = 12$ degrees of freedom.

(ii) The observed mean value of the data is 80.02. A 95% confidence interval for the true mean is $[80.006, 80.035]$.

(iii) The R functions qt(p,df) calculates the quantile from the probability density and the degrees of freedom and pt(q,df) calculates the probabilty density from the quantile and the degrees of freedom.

(iv) The **confidence interval** for measurement data consists of the values µ, for which the corresponding statistical test does not reject the null hypothesis.

```
qt(0.975,12)
[1] 2.178813
pt(2.178813,12)
[1] 0.975
qt(0.5,12)
[1] 0.0
pt(0.0,12)
[1] 0.5
```

P-Value

The p-value is the probability that the test statistic will take on a value that is at least as extreme (with respect to the alternative hypothesis) as the observed value of the statistic when the null hypothesis H_0 is true.

For the one-sided alternative hypothesis H_A: $\mu > \mu_0$, the p-value can be calculated as follows - the observed value of the statistics is $t = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\hat{\sigma}_X} = 3.1246$:

p-value$= P(T > t) = P(T > 3.1246) = 0.00439$ For the two-sided alternative hypothesis H_A: $\mu \neq \mu_0$, the p-value can be calculated as follows (the observed value of the test statistics is $t = \frac{\sqrt{n}|\bar{X}_n - \mu_0|}{\hat{\sigma}_X}$):

p-value$= 2 \cdot P(T > |t|)$

In R we compute the one-sided and the two sided p-value as follows:

```
#one-sided p-value
1-pt(3.1246, df=12)
[1] 0.004389739
#two-sided p-value
2*(1-pt(3.1246, df=12))
[1] 0.008779477
```

These p-values are evidence against the null hypothesis at 5% level. Whereas the two-sided value is statistically significant at the 5% value.

p-value and Statistical Test 1. Reject H_0 if p-value$\leqslant \alpha$ 2. Retain H_0 if p-value$> \alpha$ *The p-value is the smallest level of significance that would lead to rejection of the null hypothesis H_0 with the given data.*	The p-value quantifies how significant an alternative is: p-value ≈ 0.05 : weakly significant, "." p-value ≈ 0.01 : weakly significant, "*" p-value ≈ 0.001 : weakly significant, "**" p-value $\leqslant 10^{-4}$: weakly significant, "***"

2.3 Joint Distributions

TO DO: Chapter 3

2.3.1 Joint, Marginal and Conditional Distributions

Discrete Joint Probability Distribution	
The **Joint Probability Distribution** of X and Y is defined by the following distributions:	$P(X = x, Y = y)$, $x \in W_x, y \in W_y$
Marginal Distributions are single distributions $P(X = x)$ of X and $P(Y = y)$ of Y. They can be calculated based on their joint distribution: Joint distribution of (X, Y) starting from the marginal distribution of X and Y is only possible for **independent** X and Y. Then it holds:	$P(X = x) = \sum_{y \in W_y} P(X = x, Y = y)$, $x \in W_x$ $P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$, $x \in W_x, y \in W_y$
Conditional probability of X given $Y = y$ is defined as: The **marginal distributions** then can be expressed as follows:	$P(X = x \vert Y = y) = \frac{P(X=x,Y=y)}{P(Y=y}$ $P(X = x) = \sum_{y \in W_y} P(X = x \vert Y = y) P(Y = y)$, $x \in W_x$
Conditional Expected Value of Y given $X = x$ is defined as:	$E[Y \vert X = x] = \sum_{y \in W_y} y \cdot P(Y = y \vert X = x)$

Example

X\ Y	1	2	3	4	$\sum$
1	0.080	0.015	0.003	0.002	0.100
2	0.050	0.350	0.050	0.050	0.500
3	0.030	0.060	0.180	0.030	0.300
4	0.001	0.002	0.007	0.090	0.100
$\sum$	0.161	0.427	0.240	0.172	1

$P(X = 3, Y = 4) = 0.030$ or $P(X = 3 \cup Y = 4) = 0.030$

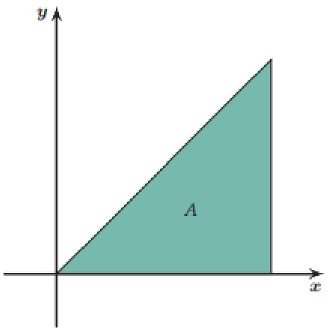
$P(X = 3) = P(X = 3, Y = 1) + P(X = 3, Y = 2)+$
$P(X = 3, Y = 3) + P(X = 3, Y = 4) =$
$0.030 + 0.060 + 0.180 + 0.030 = 0.300$

$P(Y = 2 | X = 4) = \frac{P(Y=2,X=4)}{P(X=4)} = \frac{0.002}{0.1} = 0.02$

$P(X = Y) = P(X = 1, Y = 1) + P(X = 2, Y = 2)+$
$P(X = 3, Y = 3) + P(X = 4, Y = 4) = 0.700$

If random variables are independent it must hold that
$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$
From the marginal distribution follows
$P(X = 1) \cdot P(Y = 2) = 0.100 \cdot 0.427 = 0.043$
and this is not equal to
$P(X = 1, Y = 2) = 0.15$
X and Y are **not independent**

Joint Density Function

The **probability** that the **joint random variable** (X,Y) lies in a two-dimensional region A, i.e., $A \subset \mathbb{R}^2$, is given by	$P((X,Y) \in A) = \iint\limits_A f_{X,Y}(x,y)dx\,dy$
The (bivariate) **joint density function** needs to satisfy	$\iint\limits_{\mathbb{R}} f_{X,Y}(x,y)dx\,dy = 1$
X and Y are only **independent** if	$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y),\ x,y \in \mathbb{R}$
Marginal Density	$f_X(x) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)dy,\ f_Y(y) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)dx$
Conditional Probability	$f_{Y\|X=x}(y) = f_Y(y\|X=x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$
X and Y are only independent if the following apply:	$f_{Y\|X=x}(y) = f_Y(y)$ resp. $f_{X\|Y=y}(x) = f_X(x)$
Conditional Expected Value of a continuous random variable Y given $X = x$	$E[Y\|X=x] = \int\limits_{-\infty}^{\infty} y \cdot f_{Y\|X=x}(y)dy$
Example	
Two machines with exponentially distributed life expectancy $X \sim Exp(\lambda_1)$ and $Y \sim Exp(\lambda_2)$, where X and Y are independent. $f_X(x) = \lambda_1 e^{-\lambda_1 x}$ and $f_Y(y) = \lambda_2 e^{-\lambda_2 y}$  	Due to independence: $_{X,Y}(x,y) = \lambda_1 e^{-\lambda_1 x}\lambda_2 e^{-\lambda_2 y}$ $P(Y < X) = \int\limits_0^{\infty}(\int\limits_0^{x} \lambda_1 e^{-\lambda_1 x}\lambda_2 e^{-\lambda_2 y}dy)dx$ $P(Y < X) = \int\limits_0^{\infty} \lambda_1 e^{-\lambda_1 x}(1 - e^{-\lambda_2 y})dx = \frac{\lambda_2}{\lambda_1+\lambda_2}$

2.3.2 Covariance and Correlation

Covariance and Correlation	
Covariance	$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E[XY] - E[X]E[Y]$
X,Y independent	$E[XY] = E[X]E[Y]$
	$Cov(X,X) = E[(X-\mu_X)(X-\mu_X)] = E[(X-\mu_X)^2] = Var(X)$
Sum of Variances	$Var(\sum\limits_{i=1}^{n} X_i) = Cov(\sum\limits_{i=1}^{n} X_i, \sum\limits_{i=1}^{n} X_i) = \sum\limits_{i=1}^{n} Var(X_i) + 2\sum\limits_{i<j}^{n} Cov(X_i,X_j)$
2 Random Variables	$Var(X+Y) = Cov(X+Y,X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$
If all X_i are independent	$Var(X_1 + X_2 + ... + X_n) = Var(X_1) + ... + Var(X_n)$
Correlation	$Cor(X,Y) = \rho_{XY} = \frac{Cov(X,Y}{\rho_X \rho_Y}$ where $-1 \leqslant Cor(X,Y) \leqslant 1$

Measure for strength and direction of the *linear dependency* between X and Y.	$Cor(X,Y) = +1$ if $Y = a + bX$ for $a \in \mathbb{R}$ and $b > 0$ $Cor(X,Y) = -1$ if $Y = a + bX$ for $a \in \mathbb{R}$ and $b < 0$ $\lvert Cor(X,Y)\rvert = 1$ means perfect linear relationship between X and Y. $Cor(X,Y) = 0$ means X and Y are uncorrelated.
X and Y **linear independent**	$Cor(X,Y) = 0$ (and thus $Cov(X,Y) = 0$)

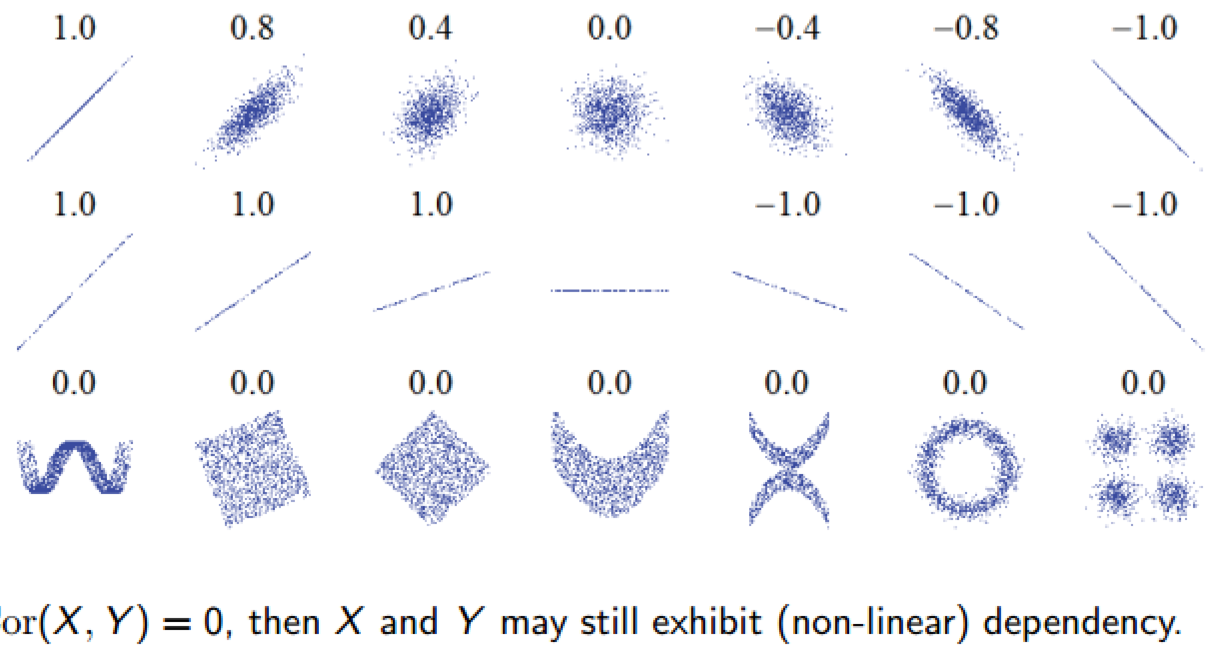


Figure 5: Correlations

2.3.3 Bivariate Normal Distribution

Bivariate Normal Distribution	
Expected values and variances of the marginal distribution Covariance between X and Y	μ_X, σ_X^2 and μ_Y, σ_Y^2 $Cov(X,Y) = \rho_{XY}\sigma_X\sigma_Y$
Joint Density	$f_{X,Y}(x,y) =$ $\frac{1}{2\pi\sqrt{det(\sum)}} exp\Big(-\frac{1}{2}(x-\mu_X, y-\mu_Y)\sum^{-1}\begin{pmatrix} x-\mu_X \\ y-\mu_Y \end{pmatrix}\Big)$
Covariance Matrix	$\sum = \begin{pmatrix} Cov(X,X) & Cov(X,Y) \\ Cov(Y,X) & Cov(Y,Y) \end{pmatrix} = \begin{pmatrix} \sigma_X^2 & \rho_{XY}\sigma_X\sigma_Y \\ \rho_{XY}\sigma_X\sigma_Y & \sigma_Y^2) \end{pmatrix}$

2.3.4 Principal Component Analysis (PCA)

3 Regression Analysis

3.1 Simple Linear Regression

TO DO: Chapter 5

...

3.1.1 Estimating the Coefficients

Theory	**Code Example**
Estimation of response variable Y based on a predictor variable X. $$Y \simeq \beta_0 + \beta_1 X$$	`lm(Y ~ X, data=someData)`

Source code:

```
advertising <- read.csv("../Data/Advertising.csv")
model <- lm(sales ~ TV, data=advertising)
summary(model)
```

Output:

```
##
## Call:
## lm(formula = sales ~ TV, data = Advertising)
##
## Residuals:
## Min 1Q Median 3Q Max
## -8.3860 -1.9545 -0.1913 2.0671 7.2124
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 7.032594 0.457843 15.36 <2e-16 ***
## TV 0.047537 0.002691 17.67 <2e-16 ***
## ---
## Signif. codes:
## 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.259 on 198 degrees of freedom
## Multiple R-squared: 0.6119,Adjusted R-squared: 0.6099
## F-statistic: 312.1 on 1 and 198 DF, p-value: < 2.2e-16
```

Interpretation of output:

TO DO: interpretation here

3.2 Residual Analysis

TO DO: Chapter 6

3.3 Multiple Linear Regression

TO DO: Chapter 7

3.4 Linear Model Selection

TO DO: Chapter 8

4 Classification

4.1 Logistic Regression

TO DO: Chapter 10

4.2 Decision Trees

Decision trees are applied to both, classification and regression. **TO DO: Chapter 11**

4.2.1 Classification Trees

4.2.1.1 Binary Splitting

In binary splitting, a training set is used to split up the predictor domain into regions which contain data for which the response variable belongs to the same class. By **binary** it is meant that a region is split into **two** subregions (i.e. "is a predictor less or greater than a threshold value?" → yes/no).

Theory

Algorithm:

1. Initialise the set of regions $\mathcal{R} = R$ by the predictor domain R
2. Choose the optimal region R in $\mathcal{R}$ and the optimal predictor X_i such that a binary split of R with respect to X

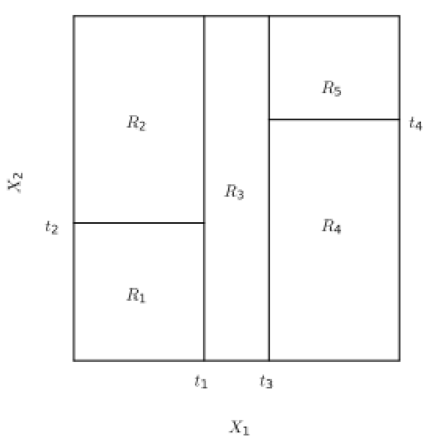
   $$R_1 = \{\vec{x} \in R | x_i > t\} \quad \text{and} \quad R_2 = \{\vec{x} \in R | x_i \leqslant t\}$$

   gives the highest gain in purity (for some threshold t).
3. Replace R in $\mathcal{R}$ with R_1 and R_2 and return to 2.

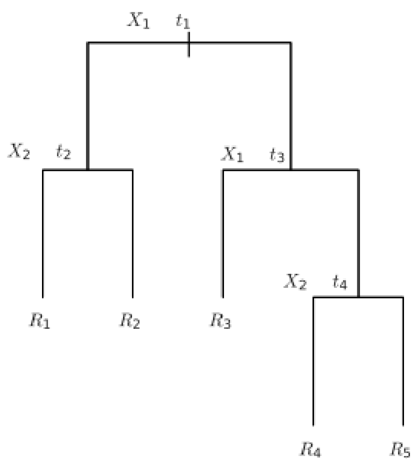
The iteration is stopped if the current splitting fulfils a predefined stopping criterion.

Code Example

```
require(tree)
#default controls
tc = tree.control(nobs = 303, mincut = 5,
    minsize = 10, mindev = 0.01)

#grow tree
tree.model = tree(AHD~MaxHR+Age, data = heart,
    control = tc)

#plot tree and label splits
plot(tree.model)
text(tree.model, cex=0.8)

#plot partition (only for two predictor case)
partition.tree(tree.model)
points(Age~MaxHR, data = heart, col = cols[label
    ], pch=20)
```



(a) Example regions resulting from binary splitting



(b) Example decision tree resulting from binary splitting

4.2.1.2 Node Purity

Notation:

Variable	Description
Y	Response variable
K	Levels (categories) of the response variable
T	The decision tree
M	Amount of terminal nodes
$\hat{p}_{mk}$	proportion of the training data in region m from level k

Purity Measures:

Classification error rate	$E_m(T) = 1 - \max_k(\hat{p}_{mk})$
Gini index	$G_m(T) = \sum_{k=1}^{K} \hat{p}_{mk} \cdot (1 - \hat{p}_{mk})$
Cross-entropy	$D_m(T) = -\sum_{k=1}^{K} \hat{p}_{mk} \cdot \log(\hat{p}_{mk})$

Code example: Cross Entropy and Gini measures in R

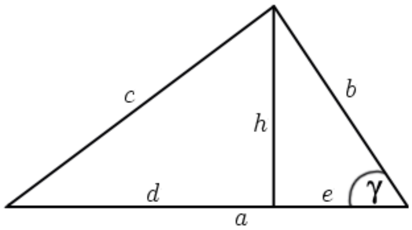
```
require(tree)
# deviance or cross entropy
tree.model = tree(AHD~MaxHR+Age, data = heart, split = "deviance")
plot(tree.model)
text(tree.model, cex=0.8)
partition.tree(tree.model)
points(Age~MaxHR, data = heart, col = cols[label], pch=20)

# Gini index
tc = tree.control(303, mincut = 5, minsize = 60, mindev = 0.01)
tree.model = tree(AHD~MaxHR+Age, data = heart, split = "gini", control = tc)
plot(tree.model)
text(tree.model, cex=0.8)
partition.tree(tree.model)
points(Age~MaxHR, data = heart, col = cols[label], pch=20)
```

5 Idiotenseite

5.1 Dreiecksformeln

Cosinussatz

$$c^2 = a^2 + b^2 - 2 \cdot a \cdot b \cdot \cos\gamma$$



Sinussatz

$$\frac{a}{\sin\alpha} = \frac{b}{\sin\beta} = \frac{c}{\sin\gamma} = 2r = \frac{u}{\pi}$$

Pythagoras beim Sinus

$$\sin^2(b) + \cos^2(b) = 1 \qquad \tan(b) = \frac{\sin(b)}{\cos(b)}$$

$$\sin\beta = \tfrac{b}{a} = \tfrac{\text{Gegenkathete}}{\text{Hypotenuse}} \qquad \tan\beta = \tfrac{c}{b} = \tfrac{\text{Gegenkathete}}{\text{Ankathete}}$$
$$\cos\beta = \tfrac{c}{a} = \tfrac{\text{Ankathete}}{\text{Hypotenuse}} \qquad \cot\beta = \tfrac{c}{b} = \tfrac{\text{Ankathete}}{\text{Gegenkathete}}$$

5.2 Funktionswerte für Winkelargumente

deg	rad	sin	cos	tan
0°	0	0	1	0
30°	$\frac{\pi}{6}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{3}$
45°	$\frac{\pi}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	1
60°	$\frac{\pi}{3}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$

deg	rad	sin	cos
90°	$\frac{\pi}{2}$	1	0
120°	$\frac{2\pi}{3}$	$\frac{\sqrt{3}}{2}$	$-\frac{1}{2}$
135°	$\frac{3\pi}{4}$	$\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$
150°	$\frac{5\pi}{6}$	$\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$

deg	rad	sin	cos
180°	π	0	-1
210°	$\frac{7\pi}{6}$	$-\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$
225°	$\frac{5\pi}{4}$	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$
240°	$\frac{4\pi}{3}$	$-\frac{\sqrt{3}}{2}$	$-\frac{1}{2}$

deg	rad	sin	cos
270°	$\frac{3\pi}{2}$	-1	0
300°	$\frac{5\pi}{3}$	$-\frac{\sqrt{3}}{2}$	$\frac{1}{2}$
315°	$\frac{7\pi}{4}$	$-\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$
330°	$\frac{11\pi}{6}$	$-\frac{1}{2}$	$\frac{\sqrt{3}}{2}$

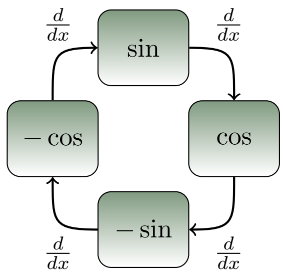
5.3 Periodizität

$\cos(a + k \cdot 2\pi) = \cos(a) \qquad \sin(a + k \cdot 2\pi) = \sin(a) \qquad (k \in \mathbb{Z})$

5.4 Quadrantenbeziehungen

$\sin(-a) = -\sin(a) \qquad \cos(-a) = \cos(a)$
$\sin(\pi - a) = \sin(a) \qquad \cos(\pi - a) = -\cos(a)$
$\sin(\pi + a) = -\sin(a) \qquad \cos(\pi + a) = -\cos(a)$
$\sin\left(\frac{\pi}{2} - a\right) = \sin\left(\frac{\pi}{2} + a\right) = \cos(a) \qquad \cos\left(\frac{\pi}{2} - a\right) = -\cos\left(\frac{\pi}{2} + a\right) = \sin(a)$

5.5 Ableitungen



5.6 Additionstheoreme

$\sin(a \pm b) = \sin(a) \cdot \cos(b) \pm \cos(a) \cdot \sin(b)$
$\cos(a \pm b) = \cos(a) \cdot \cos(b) \mp \sin(a) \cdot \sin(b)$
$$\tan(a \pm b) = \frac{\tan(a) \pm \tan(b)}{1 \mp \tan(a) \cdot \tan(b)}$$

5.7 Doppel- und Halbwinkel

$\sin(2a) = 2\sin(a)\cos(a)$
$\cos(2a) = \cos^2(a) - \sin^2(a) = 2\cos^2(a) - 1 = 1 - 2\sin^2(a)$
$\cos^2\left(\frac{a}{2}\right) = \frac{1+\cos(a)}{2} \qquad \sin^2\left(\frac{a}{2}\right) = \frac{1-\cos(a)}{2}$

5.8 Produkte

$\sin(a)\sin(b) = \frac{1}{2}(\cos(a-b) - \cos(a+b))$
$\cos(a)\cos(b) = \frac{1}{2}(\cos(a-b) + \cos(a+b))$
$\sin(a)\cos(b) = \frac{1}{2}(\sin(a-b) + \sin(a+b))$

5.9 Euler-Formeln

$\sin(x) = \frac{1}{2j}\left(e^{jx} - e^{-jx}\right) \qquad \cos(x) = \frac{1}{2}\left(e^{jx} + e^{-jx}\right)$
$e^{x+jy} = e^x \cdot e^{jy} = e^x \cdot (\cos(y) + j\sin(y))$
$e^{j\pi} = e^{-j\pi} = -1$

5.10 Summe und Differenz

$\sin(a) + \sin(b) = 2 \cdot \sin\left(\frac{a+b}{2}\right) \cdot \cos\left(\frac{a-b}{2}\right)$
$\sin(a) - \sin(b) = 2 \cdot \sin\left(\frac{a-b}{2}\right) \cdot \cos\left(\frac{a+b}{2}\right)$

$\cos(a) + \cos(b) = 2 \cdot \cos\left(\frac{a+b}{2}\right) \cdot \cos\left(\frac{a-b}{2}\right)$
$\cos(a) - \cos(b) = -2 \cdot \sin\left(\frac{a+b}{2}\right) \cdot \sin\left(\frac{a-b}{2}\right)$
$\tan(a) \pm \tan(b) = \dfrac{\sin(a \pm b)}{\cos(a)\cos(b)}$

5.12 Ableitungen elementarer Funktionen S436

Funktion	Ableitung	Funktion	Ableitung
C (Konstante)	0	$\sec x$	$\frac{\sin x}{\cos^2 x}$
x	1	$\sec^{-1} x$	$\frac{-\cos x}{\sin^2 x}$
$x^n \ (n \in \mathbb{R})$	nx^{n-1}	$\arcsin x \quad (\|x\| < 1)$	$\frac{1}{\sqrt{1-x^2}}$
$\frac{1}{x}$	$-\frac{1}{x^2}$	$\arccos x \quad (\|x\| < 1)$	$-\frac{1}{\sqrt{1-x^2}}$
$\frac{1}{x^n}$	$-\frac{n}{x^{n+1}}$	$\arctan x$	$\frac{1}{1+x^2}$
$\sqrt{x}$	$\frac{1}{2\sqrt{x}}$	$\text{arccot}\ x$	$-\frac{1}{1+x^2}$
$\sqrt[n]{x} \quad (n \in \mathbb{R}, n \neq 0, x > 0)$	$\frac{1}{n\sqrt[n]{x^{n-1}}}$	$\text{arcsec}\ x$	$\frac{1}{x\sqrt{x^2-1}}$
e^x	e^x	$\text{arcossec}\ x$	$-\frac{1}{x\sqrt{x^2-1}}$
$e^{bx} \quad (b \in \mathbb{R})$	be^{bx}	$\sinh x$	$\cosh x$
$a^x \quad (a > 0)$	$a^x \ln a$	$\cosh x$	$\sinh x$
$a^{bx} \quad (b \in \mathbb{R}, a > 0)$	$ba^{bx} \ln a$	$\tanh x$	$\frac{1}{\cosh^2 x}$
$\ln x$	$\frac{1}{x}$	$\coth x \quad (x \neq 0)$	$-\frac{1}{\sinh^2 x}$
$\log_a x \quad (a > 0, a \neq 1, x > 0)$	$\frac{1}{x}\log_a e = \frac{1}{x \ln a}$	$\text{Arsinh}\ x$	$\frac{1}{\sqrt{1+x^2}}$
$\lg x \quad (x > 0)$	$\frac{1}{x}\lg e \approx \frac{0.4343}{x}$	$\text{Arcosh}\ x \quad (x > 1)$	$\frac{1}{\sqrt{x^2-1}}$
$\sin x$	$\cos x$	$\text{Artanh}\ x \quad (\|x\| < 1)$	$\frac{1}{1-x^2}$
$\cos x$	$-\sin x$	$\text{Arcoth}\ x \quad (\|x\| > 1)$	$-\frac{1}{x^2-1}$
$\tan x \quad (x \neq (2k+1)\frac{\pi}{2}, k \in \mathbb{Z})$	$\frac{1}{\cos^2 x} = \sec^2 x$	$[f(x)]^n \quad (n \in \mathbb{R})$	$n[f(x)]^{n-1} f'(x)$
$\cot x \quad (x \neq k\pi, k \in \mathbb{Z})$	$\frac{-1}{\sin^2 x} = -cosec^2 x$	$\ln f(x) \quad (f(x) > 0)$	$\frac{f'(x)}{f(x)}$

5.11 Einige unbestimmte Integrale S1074

$\int dx = x + C$	$\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C,\ x\epsilon\mathbb{R}^+,\ \alpha\epsilon\mathbb{R}\backslash\{-1\}$
$\int \frac{1}{x}dx = \ln\lvert x\rvert + C,\ x \neq 0$	$\int e^x dx = e^x + C$
$\int a^x dx = \frac{a^x}{\ln a} + C,\ a\epsilon\mathbb{R}^+\backslash\{1\}$	$\int \sin x dx = -\cos x + C$
$\int \cos x dx = \sin x + C$	$\int \frac{dx}{\sin^2 x} = -\cot x + C,\ x \neq k\pi$ mit $k\epsilon\mathbb{Z}$
$\int \frac{dx}{\cos^2 x} = \tan x + C,\ x \neq \frac{\pi}{2} + k\pi$ mit$k\epsilon\mathbb{Z}$	$\int \sinh x dx = \cosh x + C$
$\int \cosh x dx = \sinh x + C$	$\int \frac{dx}{\sinh^2 x} = -\coth x + C,\ x \neq 0$
$\int \frac{dx}{\cosh^2 x} = \tanh x + C$	$\int \frac{dx}{ax+b} = \frac{1}{a}\ln\lvert ax+b\rvert + C,\ a \neq 0, x \neq -\frac{b}{a}$
$\int \frac{dx}{a^2x^2+b^2} = \frac{1}{ab}\arctan\frac{a}{b}x + C,\ a \neq 0,\ b \neq 0$	$\int \frac{dx}{a^2x^2-b^2} = \frac{1}{2ab}\ln\left\lvert\frac{ax-b}{ax+b}\right\rvert + C,\ a \neq 0,\ b \neq 0,\ x \neq \frac{b}{a},\ x \neq -\frac{b}{a}$
$\int \sqrt{a^2x^2+b^2}dx = \frac{x}{2}\sqrt{a^2x^2+b^2} + \frac{b^2}{2a}\ln\left(ax + \sqrt{a^2x^2+b^2}\right) + C,\ a \neq 0,\ b \neq 0$	$\int \sqrt{a^2x^2-b^2}dx = \frac{x}{2}\sqrt{a^2x^2-b^2} - \frac{b^2}{2a}\ln\left\lvert ax + \sqrt{a^2x^2-b^2}\right\rvert + C,\ a \neq 0,\ b \neq 0, a^2x^2 \geqq b^2$
$\int \sqrt{b^2-a^2x^2}dx = \frac{x}{2}\sqrt{b^2-a^2x^2} + \frac{b^2}{2a}\arcsin\frac{a}{b}x + C,\ a \neq 0,\ b \neq 0,\ a^2x^2 \leqq b^2$	$\int \frac{dx}{\sqrt{a^2x^2-b^2}} = \frac{1}{a}\ln(ax + \sqrt{a^2x^2+b^2}) + C,\ a \neq 0,\ b \neq 0$
$\int \frac{dx}{\sqrt{a^2x^2-b^2}} = \frac{1}{a}\ln\left\lvert ax + \sqrt{a^2x^2-b^2}\right\rvert + C,\ a \neq 0,\ b \neq 0,\ a^2x^2 > b^2$	$\int \frac{dx}{\sqrt{b^2-a^2x^2}} = \frac{1}{a}\arcsin\frac{a}{b}x + C,\ a \neq 0,\ b \neq 0,\ a^2x^2 < b^2$
Die Integrale $\int \frac{dx}{X}, \int \sqrt{X}dx, \int \frac{dx}{\sqrt{X}}$ mit $X = ax^2 + 2bx + c,\ a \neq 0$ werden durch die Umformung $X = a(x+\frac{b}{a})^2 + (c - \frac{b^2}{a})$ und die Substitution $t = x + \frac{b}{a}$ in die oberen 4 Zeilen transformiert.	$\int \frac{xdx}{X} = \frac{1}{2a}\ln\lvert X\rvert - \frac{b}{a}\int\frac{dx}{X},\ a \neq 0,\ X = ax^2 + 2bx + c$
$\int \sin^2 ax dx = \frac{x}{2} - \frac{1}{4a}\cdot\sin 2ax + C,\ a \neq 0$	$\int \cos^2 ax dx = \frac{x}{2} + \frac{1}{4a}\cdot\sin 2ax + C,\ a \neq 0$
$\int \sin^n ax dx = -\frac{sin^{n-1}ax\cdot\cos ax}{na} + \frac{n-1}{n}\int\sin^{n-2}ax dx,\ n\epsilon\mathbb{N},\ a \neq 0$	$\int \cos^n ax dx = \frac{\cos^{n-1}ax\cdot\sin ax}{na} + \frac{n-1}{n}\int\cos^{n-2}ax dx,\ n\epsilon\mathbb{N},\ a \neq 0$
$\int \frac{dx}{\sin ax} = \frac{1}{a}\ln\left\lvert\tan\frac{ax}{2}\right\rvert + C,\ a \neq 0,\ x \neq k\frac{\pi}{a}$ mit $k\epsilon\mathbb{Z}$	$\int \frac{dx}{\cos ax} = \frac{1}{a}\ln\left\lvert\tan(\frac{ax}{2} + \frac{\pi}{4})\right\rvert + C,\ a \neq 0,\ x \neq \frac{\pi}{2a} + k\frac{\pi}{a}$ mit $k\epsilon\mathbb{Z}$
$\int \tan ax dx = -\frac{1}{a}\ln\lvert\cos ax\rvert + C,\ a \neq 0,\ x \neq \frac{\pi}{2a} + k\frac{\pi}{a}$mit $k\epsilon\mathbb{Z}$	$\int \cot ax dx = \frac{1}{a}\ln\lvert\sin ax\rvert + C,\ a \neq 0,\ x \neq k\frac{\pi}{a}mitk\epsilon\mathbb{Z}$
$\int x^n \sin ax dx = -\frac{x^n}{a}\cos ax + \frac{n}{a}\int x^{n-1}\cos ax dx,\ n\epsilon\mathbb{N},\ a \neq 0$	$\int x^n \cos ax dx = \frac{x^n}{a}\sin ax - \frac{n}{a}\int x^{n-1}\sin ax dx,\ n\epsilon\mathbb{N},\ a \neq 0$
$\int x^n e^{ax} dx = \frac{1}{a}x^n e^{ax} - \frac{n}{a}\int x^{n-1}e^{ax}dx,\ n\epsilon\mathbb{N},\ a \neq 0$	$\int e^{ax}\sin bx dx = \frac{e^{ax}}{a^2+b^2}(a\sin bx - b\cos bx) + C,\ a \neq 0,\ b \neq 0$
$\int e^{ax}\cos bx dx = \frac{e^{ax}}{a^2+b^2}(a\cos bx + b\sin bx) + C,\ a \neq 0,\ b \neq 0$	$\int \ln x dx = x(\ln x - 1) + C,\ x\epsilon\mathbb{R}^+$
$\int x^\alpha\cdot\ln x dx = \frac{x^{\alpha+1}}{(\alpha+1)^2}[(\alpha+1)\ln x - 1] + C,\ x\epsilon\mathbb{R}^+,\ \alpha\epsilon\mathbb{R}\backslash\{-1\}$	