1 Kombinatorika

Pravilo produkta

Denimo da izbiranje poteka v k zaporednih korakih. Na prvem koraku izbiramo med n_1 moznostmi, in nato na drugem med n_2 moznostmi, ..., in nato na k-tem koraku med n_k moznostmi. Na vsakem koraku izbiranja naj bo stevilo moznosti, neodvsino od izbranih moznosti na prejsnih korakih. Stevilo vseh moznih izborov je tedaj enako:
$n_1 \cdot n_2 \cdot n_3 \cdot ... \cdot n_k$

Pravilo vsote

Kadar se pri izbiranju odlocamo za eno od n_1 moznosti iz prve mnozice, ali za eno od n_2 moznosti iz druge mnozice, ..., ali za eno od n_k moznosti iz k-te mnozice, in so te mnozice paroma disjunktne je stevilo vseh moznih izidov enako:
$n_1 + n_2 + n_3 + ... + n_k$

Pravilo komplementa

Ce je vseh moznosti n in je s takih, ki ne ustrezajo pogojem, potem je moznosti:
$d = n - s$

Permutacije brez ponavljanja

so razporeditve n razlicnih elementov na n prostih mest. Vrstni red elementov je pomemben. Stevilo vseh takih razporeditev je:
$n! = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 2 \cdot 1$

Variacije brez ponavljanja

Razporeditev n razlicnih elementov na k prostih mest, vrstni red je pomemben. $\frac{n!}{(n-k)!} = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot (n-k+1)$

Permutacije s ponavlanjem

so razporeditve n ne nujno razlicnih elementov na n prostih mest. Vrstni red je pomemben. Ce lahko elemente razdelimo v m skupin med sabo enakih elementov , in ce je v prvi taki skupini k_1 enakih elementov, v drugi pa k_2 enakih elementov, ..., v m-ti pa k_m enakih elementov, potem je stevilo vseh takih razporeditev:
$\binom{n}{k_1, k_2, ..., k_m} = \frac{n!}{k_1! \cdot k_2! \cdot ... \cdot k_m!}$

Variacije s ponavlanjem

so razporeditve n razlicnih elementov na k prostih mest. Vrstni red je pomemben. Pri tem se lahko element dolocene vrste v razporeditvi pojavi poljubno mnogokrat. Stevilo vseh takih razporeditev je:
$n^k = n \cdot n \cdot n \cdot ... \cdot n$

Kombinacije brez ponavljanja

so izbire k elemntov izmed n razlicnih elementov, kjer lahko vsak element izberemo le enkrat. Vrstni red izbir ni pomemben. Stevilo kombinacij brez ponavljanja je:
$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Pravilo kvocienta (vrtiljak)

Ce imamo n objektov, ki jih grupiramo v skupine velikosti k, za katere velja , da objektov znotraj iste skupine ne locimo, objekte iz dveh razlicnih skupin pa razlikujemo, potem je razlicnih razredov objektov $\frac{n}{k}$

Kombinacije s ponavljanjem

so izbire k elementov izmed n razlicnih elementov (vrstni red ni pomemben) , kjer lahko vsak element izberemo poljubnokrat. Stevilo kombinacij s ponavljanjem je:
$(\binom{n}{k}) = \binom{n+k-1}{k}$

2 Verjetnost

2.1 Elementarna verjetnost

$\Omega \equiv$ elementarna verjetnost
Izidi morajo biti enako vrjetni!
$A \subseteq \Omega \equiv$ dogodek
$P(A) \in [0,1]$
$P(\emptyset) = 0$
$P(\Omega) = 1$
$P(A) = \frac{|A|}{|\Omega|}$
Uporabne zveze:
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$P(A \cap B) = P(A) + P(B) - P(A \cup B)$
$P(A^c) = 1 - P(A)$

2.2 Geometrijska verjetnost

$P(A) = \frac{\text{mera } A}{\text{mera } B}$
Mera: dolzina, ploscina, volumen

2.3 Neovdisni dogodki

Ce sta dogodka A in B neodvisna:
$P(A \cap B) = P(A) \cdot P(B)$

2.4 Nacelo vkljucitev in izkljucitev

$P(A_1 \cup ... \cup A_2) = P(A_1) + ... + P(A_n)$
$- P(A_1A_2) - P(A_1A_3) - ... - P(A_{n-1}A_n)$
$+ P(A_1A_2A_3) + P(A_1A_2A_4) + ... + P(A_{n-2}A_{n-1}A_n)$
$- ...$
$+ (-1)^{n+1} P(A_1 ... A_n)$

$P(A \cup B \cup C) =$
$P(A) + P(B) + P(C)$
$- P(A \cap B) - P(A \cap C) - P(B \cap C)$
$+ P(A \cap B \cap C)$

3 Pogojna Verjetnost

3.1 Pogojna verjetnost

$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$
$P(A \cap B) = P(A \mid B)P(B)$

3.2 Popolna verjetnost

H_1 ... H_n so popolen sistem dogodkov ce:
$\Omega = H_1 \cup H_2 \cup ... \cup H_n$ in $H_i \cap H_j = \emptyset$
Sledi:
$P(A) = P(A \cap H_1) + ... + P(A \cap H_n)$
$P(A) = P(A|H_1)P(H_1) + ... + P(A|H_n)P(H_n)$
$P(A \mid H_2) = 1 - P(\overline{A} \mid H_2)$

3.3 Unija Hipotez

$P(A \mid H_1 \cup H_2) = \frac{P(A|H_1)P(H_1) + P(A|H_2)P(H_2)}{P(H_1) + P(H_2)}$

3.4 Bayesova formula

$P(H_i \mid A) = \frac{P(H_i)P(A|H_i)}{P(A)} =$
$= \frac{P(H_i)P(A|H_i)}{\sum_{k=1}^{n} P(H_k)P(A|H_k)}$

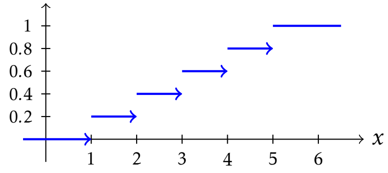
4 Diskretne slucajne spremenljivke in porazdelitve

4.1 Diskretne slucajne spremenljivke

X je realna funkcija s koncno ali **stevno** zalogo vrednosti Z_X
Porazdelitvena tabela:
$Z_X = \{x_1, x_2, ...\}$
$X \sim \begin{pmatrix} x_1 & x_2 & x_3 & \dots \\ p_1 & p_2 & p_3 & \dots \end{pmatrix}$
$P(X = x_i) = p_i$

$0 \le p_i \le 1$ in $\Sigma p_i = 1$

Porazdelitvena funkcija:
$F_X : \mathbb{R} \to [0,1]$
$F_X(x) = P(X \le x)$



$\lim_{x\to\infty} F_X(x) = 1$
$\lim_{x\to-\infty} F_X(x) = 0$
F_x je narascujoca
$\lim_{y\downarrow x} F_X(y) = F_X(x)$

4.2 Bernoullijeva

$X \sim B(p)$

- V vsakem poskusu ima dogodek A verjetnost p, X ima vrednost 1, ce se je zgodil dogodek A, 0 sicer.
- $P(X=1) = p, P(X=0) = 1-p$
- posebni primer Binomske za $n=1$

4.3 Binomska

$X \sim B(n,p)$

- X je število pojavitev izida A v n ponovitvah poskusa.
  X_2=st. lihih izidov pri 4 metih
- $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$
  za $k = 0, 1, \ldots, n$.
- R: $P(X=k) = \text{dbinom}(k,n,p), F_X(k) = \text{pbinom}(k,n,p)$.

4.4 Geometrijska

$X \sim G(p)$

- X je število ponovitev poskusa do (vključno) prve pojavitve izida A X_1=st. poskusov da pade prvic 6
- $P(X=k) = (1-p)^{k-1} p$
  $P(X \le k) = 1-(1-p)^k$
  $k = 1, 2, \ldots$
- R: $P(X=k) = \text{dgeom}(k-1,p)$
  $F_X(k) = \text{pgeom}(k-1,p)$.

4.5 Pascalova/Negativna binomska

$X \sim P(n,p)$

- X je število ponovitev poskusa do (vključno) n-te pojavitve izida A.
  Npr. koliko poskusov rabimo do 10. sestice
- $P(X=k) = \binom{k-1}{n-1}(1-p)^{k-n} p^n$
  $k = n, n+1, n+2, \ldots$
- R: $P(X=k) = \text{dnbinom}(k-n,n,p)$
  $F_X(k) = \text{pnbinom}(k-n,n,p)$.

4.6 Hipergeometrijska

$X \sim H(R,B,n)$

- X je število rdečih kroglic med izbranimi n kroglicami
- $P(X=k) = \dfrac{\binom{R}{k}\binom{B}{n-k}}{\binom{R+B}{n}}$
  $k = 0, 1, 2, \ldots, \min\{n, R\}$.
- R: $P(X=k) = \text{dhyper}(k,R,B,n)$
  $F_X(k) = \text{phyper}(k,R,B,n)$.

4.7 Poissonova

$X \sim P(\lambda)$

- V povprecju imamo na intervalu λ ponovitev dogodka A
- X pa je število ponovitev dogodka A na danem intervalu.
- $P(X=k) = \dfrac{\lambda^k}{k!} \cdot e^{-\lambda}$ za $k = 0, 1, 2, \ldots$
- R: $P(X=k) = \text{dpois}(k,\lambda)$
  $F_X(k) = \text{ppois}(k,\lambda)$.

5 Zvezne slucajne spremenljivke

5.1 Zvezna slucajna spremenljivka

Realna funkcija, za katero obstaja integrabilna funkcija $p_X : \mathbb{R} \to [0,\infty)$
$F_X(x) = P(X \le x) = \int_{-\infty}^{x} p_X(t)dt$
$p_X(x)$ funkcija gostote verjetnosti
$F_X(x)$ porazdelitvena funkcija
Z_X zaloga vrednosti (vse mozne X)
$P(a < X < b) = \int_a^b p_X(x)dx$
$= F_X(b) - F_X(a)$
$\int_{-\infty}^{\infty} p_X(x)dx = 1$

5.2 Enakomerna zvezna

$X \sim U[a,b]$

- V poskusu so vsi izidi na intervalu $[a,b]$ enako verjetni.
- $p_X(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b], \\ 0, & \text{sicer}, \end{cases}$
  $F_X(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & x \in [a,b], \\ 1, & x > b. \end{cases}$
- R: $p_X(x) = \text{dunif}(x,a,b)$
  $F_X(x) = \text{punif}(x,a,b)$.

5.3 Eksponentna

$X \sim Exp(\lambda)$

- V povprečju imamo na časovno enoto λ ponovitev dogodka A, X pa je čas med dvema zaporednima dogodkoma.
- $p_X(x) = \begin{cases} 0, & x < 0, \\ \lambda \cdot e^{-\lambda x}, & x \ge 0, \end{cases}$
- $F_X(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\lambda x}, & x \ge 0. \end{cases}$
- R: $p_X(x) = \text{dexp}(x,\lambda)$
  $F_X(x) = \text{pexp}(x,\lambda)$.

5.4 Normalna

$X \sim N(\mu,\sigma)$

- $p_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ za $x \in \mathbb{R}$.
- V primeru $\mu = 0$ in $\sigma = 1$ dobimo *standardno normalno porazdelitev*.
- $P(a \le X \le b) = F_X(b) - F_X(a) = \phi(\frac{b-\mu}{\sigma}) - \phi(\frac{a-\mu}{\sigma})$
- $F(-x) = 1 - F(x)$
- $P(\mu - \sigma \le X \le \mu + \sigma) = 0.683$
- $P(\mu - 2\sigma \le X \le \mu + 2\sigma) = 0.954$
- $P(\mu - 3\sigma \le X \le \mu + 3\sigma) = 0.997$
- Lihi momenti $E(X^3) = E(X^5) = \cdots = 0$
- R: $p_X(x) = \text{dnorm}(x,\mu,\sigma)$
  $F_X(x) = \text{pnorm}(x,\mu,\sigma)$.

5.5 Gamma

$X \sim \Gamma(n,\lambda)$

- V povprečju imamo na časovno enoto λ ponovitev dogodka A, X pa je čas med prvo in $(n+1)$. ponovitvijo dogodka A.
- $p_X(x) = \begin{cases} 0, & x \le 0, \\ \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)}, & x > 0. \end{cases}$
- R: $p_X(x) = \text{dgamma}(x,n,\lambda)$
  $F_X(x) = \text{pgamma}(x,n,\lambda)$.

5.6 Hi kvadrat

$X \sim \chi^2(n) = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$

- Je vsota kvadratov n neodvisnih standardnih normalnih slučajnih spremenljivk.
- $p_X(x) = \begin{cases} 0, & x \le 0, \\ \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \cdot \Gamma\left(\frac{n}{2}\right)}, & x > 0. \end{cases}$
- R: $p_X(x) = \text{dchisq}(x,n), F_X(x) = \text{pchisq}(x,n)$.

6 Sredine

6.1 Matematicno upanje

Diskretna spremenljivka
$E(X) = \sum\limits_{x \in Z_x} x \cdot P(X = x)$
Zvezna spremenljivka
$E(X) = \int\limits_{-\infty}^{\infty} x \cdot p_X(x) dx$

6.2 Matematicno upanje funkcije

Diskretna:
$E(f(X)) = \sum\limits_{k=0}^{\infty} f(x_k)$
Zvezna
$E(f(X)) = \int\limits_{-\infty}^{\infty} f(x) \cdot p_X(x) dx$

6.3 Matematicna upanja standardnih spremenljivk

Porazdelitev	$E(X)$
$B(p)$	p
$B(n,p)$	$n \cdot p$
$G(p)$	$\frac{1}{p}$
$P(n,p)$	$\frac{n}{p}$
$H(R,B,n)$	$\frac{nR}{R+B}$
$P(\lambda)$	λ
$E([a,b])$	$\frac{a+b}{2}$
$Exp(\lambda)$	$\frac{1}{\lambda}$
$N(\mu,\sigma)$	μ
$\chi^2(n)$	n

6.4 Linearnost slucajnih spremenljivk

Za vsaki slucajni spremenljivki X in Y (lahko sta odvisni, lahko je ena zvezna in druga diskretna) ter $a, b \in \mathbb{R}$ velja:
$E(aX + bY) = aE(X) + bE(Y)$
$E(4X + 2) = 4E(X) + E(2)$
Ce sta X in Y nedovisni $\Rightarrow E(XY) = E(X)E(Y)$

6.5 Disperzija in standardni odklon

Disperzija-Varianca
$D(X) = E((X - E(X))^2) = E(X^2) - E(X)^2$
$D(aX + b) = a^2 D(X)$
$D(X + Y) = D(X) + D(Y) + 2Cov(X, Y)$
Neodvisne spremenljivke:
$D(X + Y) = D(X) + D(Y)$
$D(X - Y) = D(X + (-Y)) = D(X) + D(Y)$
Standardni odklon
$\sigma(X) = \sqrt{D(X)}$

6.6 Disperzije standardnih spremenljivk

Porazdelitev	$D(X)$
$B(p)$	$p(1-p)$
$B(n,p)$	$n \cdot p(1-p)$
$G(p)$	$\frac{1-p}{p^2}$
$P(n,p)$	$\frac{n \cdot (1-p)}{p^2}$
$H(R,B,n)$	$\frac{nRB \cdot (R+B-n)}{(R+B)^2(R+B-1)}$
$P(\lambda)$	λ
$D([a,b])$	$\frac{(b-a)^2}{12}$
$Exp(\lambda)$	$\frac{1}{\lambda^2}$
$N(\mu,\sigma)$	σ^2
$\chi^2(n)$	$2n$

7 Vektorji

7.1 Diskretni vekotor

Porazdelitev (X, Y) lahko podamo na dva enakovredna načina, in sicer
1. Porazdelitveno tabelo

$$(X,Y) \sim \begin{array}{c|ccccc|c} X \backslash Y & y_1 & y_2 & \dots & y_m & \dots & X \\ \hline x_1 & p_{11} & p_{12} & \dots & p_{1m} & \dots & p_1 \\ x_2 & p_{21} & p_{22} & \dots & p_{2m} & \dots & p_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & p_{n1} & p_{n2} & \dots & p_{nm} & \dots & p_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hline Y & q_1 & q_2 & \dots & q_m & \dots & 1 \end{array}$$

pri čemer je $0 \le p_{ij} \le 1$, $\sum\limits_{i=1}^{\infty}\sum\limits_{j=1}^{\infty} p_{ij} = 1$, $\sum\limits_{j=1}^{\infty} p_{ij} = p_i$ za vsak $i \in \mathbb{N}$ in $\sum\limits_{i=1}^{\infty} p_{ij} = q_j$ za vsak $j \in \mathbb{N}$
2. Porazdelitveno funkcijo
$F_{X,Y}(x,y) = P(X \le x, Y \le y)$.
$F_X(x,y) = \sum\limits_{i=1}^{\infty}\sum\limits_{j=1}^{\infty} p_{ij} \cdot I_{[x_i,\infty)}(x) I_{[y_j,\infty)}(y)$
$I_{[x_i,\infty)}(x) = \begin{cases} 1, & x_i \le x, \\ 0, & \text{sicer}, \end{cases}$
$I_{[y_j,\infty)}(y) = \begin{cases} 1, & y_j \le y, \\ 0, & \text{sicer}. \end{cases}$
Robne porazdelitve
$P(X = x_i) = \sum\limits_{j=1}^{\infty} p_{ij}$
$P(Y = y_i) = \sum\limits_{i=1}^{\infty} p_{ij}$

7.2 Neodvisnost spremenljivk X in Y

Ce za ∀ x,y velja:
$P(X = x, Y = y) = P(X = x)P(Y = y)$

7.3 Zvezni vektor

Naj bosta X, Y zvezni slučajni spremenljivki. Par (X, Y) je **zvezni slučajni vektor**, če obstaja integrabilna funkcija $p_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$, tako da za vsak par $(x, y) \in \mathbb{R}^2$ velja

$$F_{X,Y}(x,y) = P(X \le x, Y \le y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p_{X,Y}(x,y)\,dx\,dy.$$

Funkciji $p_{X,Y}$ pravimo **(dvorazsežna) gostota verjetnosti**, funkciji $F_{X,Y}$ pa **porazdelitvena funkcija**. Velja

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dx\,dy = 1.$$

Robni gostoti sta

$$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dy \quad \text{in} \quad p_Y(y) = \int_{-\infty}^{\infty} p_{X,Y}(x,y$$

7.4 Neodvisnost spremenljivk X in Y

Ce za ∀ x,y $X \le x$ in $Y \le y$
$p_{X,Y}(x,y) = p_X(x) p_Y(y)$

7.5 Matematicno upanje vektorja

diskretni slucajni vektor
$E(f(X,Y)) = \sum\limits_{i=1}^{\infty}\sum\limits_{j=1}^{\infty} f(x_i, y_j) P(X = x_i, Y = y_i)$
zvezni slucajni vektor
$E(f(X,Y)) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x,y) p_{X,Y}(x,y) dx dy$

7.6 neodivsnot

Ce sta X,Y neodvisni
$E(XY) = E(X)E(Y)$

7.7 Kovarianca slucajnih spremenljivk

$\mathrm{Cov}(X,Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$
$\mathrm{Cov}(X,Y) = 0 \Leftrightarrow X$ neodvisen od Y
$\mathrm{Cov}(aX, bY) = ab\mathrm{Cov}(X,Y)$
$\mathrm{Cov}(X + a, Y + b) = \mathrm{Cov}(X,Y)$
$\mathrm{Cov}(aX + bY, Z) = a\mathrm{Cov}(X,Z) + b\mathrm{Cov}(Y,Z)$
$\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$
$|\mathrm{Cov}(X,Y)| \le \sqrt{D(X)D(Y)}$

7.8 Korelacijski koeficient

$r(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$
$r(X,Y) = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E(X)^2} \cdot \sqrt{E(Y^2) - E(Y)^2}}$
$-1 \le r(X,Y) \le 1$
$\forall a,b,c,d \in \mathbb{R} \wedge a,c \ge 0$
$r(aX + b, cY + d) = r(X,Y)$
X in Y sta **nekorelirani** $\Leftrightarrow r(X,Y) = 0$
X in Y v **linearni zvezi** $\Leftrightarrow r(X,Y) = \pm 1$

8 CLI

8.1 Vsota normalnih spremenljivk

Naj bodo:
$X_1 \sim N(\mu_1, \sigma_1)$, $X_2 \sim N(\mu_2, \sigma_2)$, …, $X_n \sim N(\mu_n, \sigma_n)$ **neodvisne**. Potem velja:

$$X_1 + X_2 \ldots + X_n \sim N\left(\sum_{i=1}^{n} \mu_i, \sqrt{\sum_{i=1}^{n} \sigma_i^2}\right).$$

8.2 Centralni limitni izrek

Naj bodo X_1, … X_n neovidsne in enako porazdeljene slucajne spremenljivke s pricakovoano vrednostjo μ in standardnim odklonom σ. Potem za dovolj velik n velja, da je porazdelitev vsote $S = X_1 + X_2 + \ldots X_n$ priblizno normalna:

$$S \sim N(n \cdot \mu, \sigma \cdot \sqrt{n})$$

Pri aproksimaciji diskretne porazdelitve vsote uporabljamo popravek za zvezonst 0.5
$P(a \leq S \leq b) \approx P(a - 0.5 \leq Y \leq b + 0.5) = F(b + 0.5) - F(a - 0.5)$

8.3 Normalno porazdeljen vzorec

Enostavni slucajni vzorec je slucajni vekotor $(X_1, ..., X_n)$ za katerega velja:

- vsi X_i imajo enako porazdelitev
- X_i so med sabo neodvisni

Vzorcno povprecje normalno porazdeljenega vzorca Naj bo enostavni $(X_1, ..., X_n)$ slucajni vzorec, kjer $X_i \sim N(\mu, \sigma)$. Potem je:
$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ tudi normalna

$$\overline{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$$

8.4 Centralni limitni izrek za vzorncno povprecje

Naj bo enostavni $(X_1, ..., X_n)$ slucajni vzorec in:
$E(X_i) = \mu$ ter $D(X_i) = \sigma^2 < \infty$
potem za dovolj veliki vzorec $n \geq 30$:

$$\overline{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$$

9 Sestavljene slucajne spremenljivke

9.1 Nova spremenljivka kot funkcija druge

Spremenimo samo zalogo vrednosti.
$N = f(X)$
$Z_N = f(Z_X)$
$P(Z = f(x_i)) = P(X = x_i)$

9.2 Nova spremenljivka kot sestavljena iz drugih

$N = f(X_1, \ldots, X_n)$ Zaloga vrednosti zajema vse moznosti.
Znamo samo za diskretne (pregledamo iz tabelce)

10 Statistika

10.1 Kvantil

$\alpha \in [0,1], \alpha - kvantil:$
$P(X \leq q_\alpha) = \alpha$
$q_\alpha = F_X^{-1}(\alpha)$
$q_{\frac{1}{4}} \ldots$ 1. kvartil
$q_{\frac{2}{4}} \ldots$ 2. kvartil … mediana
$q_{\frac{3}{4}} \ldots$ 3. kvartil

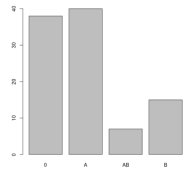
10.2 Kategoricne slucajne spremenljivke

Za vrednost imajo kategorije.
Nominalne (imenske)
med kategorijami ni hiearhije:
spol: $0 = $ moski, $1 = $ zenski
barva las: $1 = crna, 2 = rjava, 3 = rdeca, 4 = blondna, 5 = bela$
Ordinalne (urejenostne)
kategorije so urejen, obstaja hiearhija med njimi
stopnja bolecine: $0 = $ ni, $1 = $ blage, $2 = $ srednje, $3 = $ mocne
stopnja izobrazbe: $1 = $ brez, $2 = $ osnovna sola, $3 = $ srednja sola, $4 = $ fakulteta, $5 = $ postdiplomski studij

10.3 Nominalna slucajna spremenljivka

Zabelezili smo krvno skupino nakljucnega vzorca 100 Slovencev.

krvna skupina	0	A	B	AB
frekvenca f	38	40	15	7



Lahko ji izracunamo modus (najvisji stolpec)

```
krv.skupina<-c(rep("0",38),rep("A",40),
rep("B",15), rep("AB",7))
table(krv.skupina)
barplot(table(krv.skupina))
```

10.4 Ordinalna slucajna spremenljivka

10.4.1 Opisna statistika

Podatki se razvrstijo v narascujocem vrstnem redu:
$Y_1 \leq Y_2 \leq \ldots Y_n$ (00111122222233).

- Modus
- Mediana $M_e = \begin{cases} Y_{(n+1)/2}, & n - \text{liho}, \\ \frac{Y_{n/2} + Y_{n/2+1}}{2}, & n - \text{sodo}, \end{cases}$
- Kvartili
  Q_1 je mediana prve polovice sortiranih podatkov.
  Q_3 je mediana druge polovice sortiranih podatkov.

10.4.2 Grafi

- Stolpcni diagram

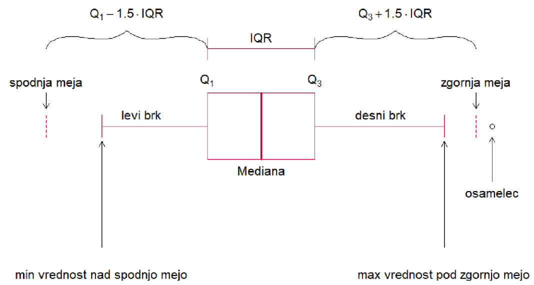
10.5 Zvezne slucajne spremenljivke

10.5.1 Opisna statistika

- **Vzorcno povprecje** $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$
  R: **mean(teza)**
- **Popravljeni vzorcni standardni odklon**
  $S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2}$
  R: **sd(teza)**
- **Povzetek s petimi stevili**
  minimum, maksimum, mediana, prvi in tretji kvartil R: **summary(teza)**

10.5.2 Grafi

- **Histrogram**
  - tezo razdelimo v razrede (intervale) in presetejemo stevilo podatkov v vsakem razredu.
  - za n podatkov je optimalno stevilo rezredov $k = \text{ceiling}(\log_2(n)) + 1$
  - dolzina razreda: $h = \frac{max - min}{k}$
- **Skatla z brki**
  - Mediana: navpicna crta v skatli
  - Meji skatle Q_1 in Q_3
  - Dolzina skatle $IQR = Q_3 - Q_1$
  - Spodnja meja $= Q_1 - 1.5 \cdot \text{IQR}$
  - Zgornja meja $= Q_1 + 1.5 \cdot \text{IQR}$
  - Podatki, ki so manjsi od spodnje ali vecji od zgornje meje se imenujejo osamelci (outliers)



10.6 Diskretne slucajne spremelnjivke

10.6.1 Opisna statistika

- minimum
- maksimum
- modus
- mediana
- kvartili
- povprecje
- standardni odklon

10.6.2 Grafi

- stolpicni diagram (malo stevilo vrednosti):
- histrogram, skatla z brki (vecje stevilo vrednosti):

11 Intervali zaupanja

11.1 Interval zaupanja za neznani parameter porazdelitve

Naj bo θ neznani parameter porazdelitve slučajne spremenljivke X in $(X_1, X_2, \ldots, X_n)$ enostavni slučajni vzorec. Iščemo interval vrednosti, v katerem se, z veliko verjetnostjo, nahaja neznani parameter θ. Na osnovi vzorca se definirata statistiki (funkciji vzorca) L in U tako da velja

$$P(L \leq \theta \leq U) = 1 - \alpha.$$

Potem rečemo, da je $I_\theta = [L,\ U]$ interval zaupanja za neznani parameter θ s *stopnjo zaupanja* $1 - \alpha$. Število α se imenuje *stopnja tveganja*.

Običajno se računa 90%, 95% ali 99% interval zaupanja ($1 - \alpha$ je enako 0.9, 0.95 ali 0.99).

11.2 Interval zaupanja za pricakovano vrednost μ

11.3 Opomba

11.4 Interval zaupanja za standardni odklon σ (pricakovana vrednost μ ni znana)

11.5 Interval zaupanja za delez p

12 Testiranje domnev

12.1 Statisticni test za en vzorec

12.2 Statisticni test za dva vzorca

12.3 Pearsonov χ^2 test

13 Intervali zaupanja

Naj bo $(X_1, X_2, ..., X_n)$ prost slucajni vzorec.
Za $(x_1, x_2, ..., x_n)$ oznacimo dobljeni vzorec konkretnih vrednosti.

Parameter	Cenilka $f(X_1,..,X_n)$	Ocena $f(x_1,..x_n)$
μ	$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
σ	$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$	$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}$
p	$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

13.1 Interval zaupanja za μ (σ je znan)

Za konstrukcijo intervala zaupanja uporabljamo dejstvo

$$\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

Slucajna spremenljivka X je normalno porazdeljena ali imamo dovolj veliki vzorec (za uporabo CLI).
Z verjetnostjo $1 - \alpha$ se μ nahaja v intervalu

$$[\overline{X} - \epsilon, \overline{X} + \epsilon]$$

Dobimo

$$\epsilon = c\frac{\sigma}{\sqrt{n}}$$

, kjer je $\underline{c = F^{-1}(1 - \frac{\alpha}{2})}$ (kvantil standardne normalne porazdelitve)
Interval zaupanja za μ

$$I_\mu = \left[\overline{X} - c\frac{\sigma}{\sqrt{n}}, \overline{X} + c\frac{\sigma}{\sqrt{n}}\right]$$

13.2 Interval zaupanja za μ (σ ni znan)

Za konstrukcijo intervala zaupanja uporabimo dejstvo:

$$\frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

Slucajna spremenljivka X je normalno porazdeljena ali imamo dovolj veliki vzorec.
Interval zaupanja za μ stopnjo zaupanja $1-\alpha$ je enak

$$I_\mu = \left[\overline{X} - c\frac{S}{\sqrt{n}}, \overline{X} + c\frac{S}{\sqrt{n}}\right]$$

kjer je $c = t_{n-1;1-\frac{\alpha}{2}}$ kvantil Studentovo porazdelitve z n-1 prostosnimi stopnjami.

13.3 Interval zaupanja za delez p

Za konstrukcijo intervala zaupanja uporabljamo dejstvo:

$$\hat{p} \sim N(p, \sqrt{\frac{p(1-p)}{n}})$$

p je delez populacije z doloceno lastnostjo.
Naj bo $(X_1, ..., X_n)$ enostavni slucajni vzorec, kjer je X_i indikatorska slucajna spremenljivka.

$$X_i \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

Neznani delez p ocenjujemo z vzorcnim delezom.

$$\hat{p} \sim \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$I_p = \left[\hat{p} - c\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + c\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

$c = F^{-1}(1 - \frac{\alpha}{2})$

14 Testiranje domneve O μ

Naj bo $(X_1, ..., X_n)$ enostavni slucajni vzorec.
Testiramo $H_0 : \mu = \mu_0$ (μ_0 je predpostavljena vrednost) proti eni od alternativnih domnev
Predpostavka: slucajna spremenljivka X je normalno porazdeljena ali je vzorec dovolj velik (za uporabo CLI).

1. $H_1 : \mu \neq \mu_0$
2. $H_1 : \mu < \mu_0$
3. $H_1 : \mu > \mu_0$

Uporabimo testno statistiko

$$Z = \frac{\overline{X} - \mu_0}{\sigma}\sqrt{n}$$

$Z \sim N(0,1)$
Kriticno obmocje K

H_1	K	c
$\mu \neq \mu_0$	$\lvert Z\rvert \geq c$	$F^{-1}(1 - \frac{\alpha}{2})$
$\mu < \mu_0$	$\lvert Z\rvert \leq c$	$F^{-1}(\alpha) = -F^{-1}(1 - \alpha)$
$\mu > \mu_0$	$\lvert Z\rvert \geq c$	$F^{-1}(1 - \alpha)$

The standard Lorem Ipsum passage, used since the 1500s

"Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum." Section 1.10.32 of "de Finibus Bonorum et Malorum", written by Cicero in 45 BC

"Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur? Quis autem vel eum iure reprehenderit qui in ea voluptate velit esse quam nihil molestiae consequatur, vel illum qui dolorem eum fugiat quo voluptas nulla pariatur?" 1914 translation by H. Rackham

"But I must explain to you how all this mistaken idea of denouncing pleasure and praising pain was born and I will give you a complete account of the system, and expound the actual teachings of the great explorer of the truth, the master-builder of human happiness. No one rejects, dislikes, or avoids pleasure itself, because it is pleasure, but because those who do not know how to pursue pleasure rationally encounter consequences that are extremely painful. Nor again is there anyone who loves or pursues or desires to obtain pain of itself, because it is pain, but because occasionally circumstances occur in which toil and pain can procure him some great pleasure. To take a trivial example, which of us ever undertakes laborious physical exercise, except to obtain some advantage from it? But who has any right to find fault with a man who chooses to enjoy a pleasure that has no annoying consequences, or one who avoids a pain that produces no resultant pleasure?" Section 1.10.33 of "de Finibus Bonorum et Malorum", written by Cicero in 45 BC

"At vero eos et accusamus et iusto odio dignissimos ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga. Et harum quidem rerum facilis est et expedita distinctio. Nam libero tempore, cum soluta nobis est eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus saepe eveniet ut et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum hic tenetur a sapiente delectus, ut aut reiciendis voluptatibus maiores alias consequatur aut perferendis doloribus asperiores repellat." 1914 translation by H. Rackham

"On the other hand, we denounce with righteous indignation and dislike men who are so beguiled and demoralized by the charms of pleasure of the moment, so blinded by desire, that they cannot foresee the pain and trouble that are bound to ensue; and equal blame belongs to those who fail in their duty through weakness of will, which is the same as saying through shrinking from toil and pain. These cases are perfectly simple and easy to distinguish. In a free hour, when our power of choice is untrammelled and when nothing prevents our being able to do what we like best, every pleasure is to be welcomed and every pain avoided. But in certain circumstances and owing to the claims of duty or the obligations of business it will frequently occur that pleasures have to be repudiated and annoyances accepted. The wise man therefore always holds in these matters to this principle of selection: he rejects pleasures to secure other greater pleasures, or else he endures pains to avoid worse pains."