Formule za 1. kolokvij iz verjetnosti

Kombinatorika

- *permutacija brez ponavljanja* je urejen razpored dolžine n, ki vsebuje n različnih elementov. Število permutacij je $n!$.
- *permutacija s ponavljanjem* je urejen razpored dolžine n, ki vsebuje k različnih elementov, vsakega $n_1, n_2, ..., n_k$, kjer je $n = n_1 + n_2 + ... + n_k$. Število le-teh je $\frac{n!}{n_1!n_2!...n_k!}$.
- *variacija brez ponavljanja* je urejen razpored dolžine k iz množice z n elementi. Število variacij je $\frac{n!}{(n-k)!}$.
- *variacija s ponavljanjem* je urejen razpored dolžine k iz množice z n elementi, ki se lahko ponavljajo. Število le-teh je n^k.
- *kombinacija brez ponavljanja* je neurejena izbira k elementov iz množice z n elementi. Število kombinacij je $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k}$.
- *kombinacija s ponavljanjem* je neurejena izbira k elementov iz množice z n elementi, ki se lahko ponavljajo. Število le-teh je $\binom{n+k-1}{k}$.

Elementarna verjetnost
Naj ima poskus n možnih izidov izmed katerih naj bo m število ugodnih izidov za dogodek A. Če so vsi izidi enakoverjetni, velja

$$P(A) = \frac{m}{n} = \frac{\text{število ugodnih izidov}}{\text{število vseh izidov}}.$$

Geometrijska verjetnost

$$P(A) = \frac{\text{mera ugodnih izidov}}{\text{mera vseh izidov}},$$

pri tem je mera lahko dolžina, ploščina, prostornina ...

Nekatere formule:

- N-nemogoč, G-gotov, $\bar{A}$-nasproten dogodek

$$P(N) = 0, \quad P(G) = 1, \quad P(\bar{A}) = 1 - P(A)$$

- A, B sta nezdružljiva dogodka, če velja $AB = N$, potem je $A \cup B = A + B$
- $P(A \cup B) = P(A) + P(B) - P(AB)$
- $P(AB) = P(A)P(B|A) = P(B)P(A|B)$
- *Bernoullijeva formula.* Če izvedemo n neodvisnih poskusov, od katerih vsak uspe z verjetnostjo p, je verjetnost, da uspe natanko k poskusov, enaka $\binom{n}{k}p^k(1-p)^{n-k}$.

Pogojna verjetnost

$$P(A|B) = \frac{P(AB)}{P(B)}$$

- Dogodka A in B sta neodvisna, če velja $P(AB) = P(A)P(B)$ oz. $P(A|B) = P(A)$.
- Če $H_1, H_2, ..., H_n$ tvorijo popoln sistem dogodkov (vedno se zgodi natanko eden izmed njih), velja *izrek o popolni verjetnosti*

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + ... + P(H_n)P(A|H_n)$$

  in *Bayesov obrazec*

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + ... + P(H_n)P(A|H_n)}.$$

Diskretne slučajne spremenljivke

Slučajna spremenljivka X je diskretna, če je njena zaloga števna množica $\{x_1, x_2, x_3, ...\}$. Porazdelitev slučajne spremenljivke X opišemo

- z verjetnostno funkcijo:

$$X \sim \left(\begin{array}{cccc} x_1 & x_2 & x_3 & \cdots \\ p_1 & p_2 & p_3 & \cdots \end{array} \right), \quad p_1 + p_2 + p_3 + \cdots = 1,$$

  kjer je $P[X = x_n] = p_n$ za $n = 1, 2, 3, ...$

- s porazdelitveno funkcijo:

$$F_X(x) = P[X < x] = \sum_{x_i < x} p_i,$$

  kjer velja $\lim_{x\to\infty} F_X(x) = 1$, $\lim_{x\to-\infty} F_X(x) = 0$, in $P[X = x_0] = \lim_{x\downarrow x_0} F_X(x) - F_X(x_0)$.

Matematično upanje diskretne spremenljivke X je

$$E(X) = p_1x_1 + p_2x_2 + p_3x_3 + \cdots .$$

Disperzija diskretne spremenljivke X je

$$D(X) = E\left((X - E(X))^2\right) = E\left(X^2\right) - E(X)^2 = \sum_n p_nx_n - \sum_n p_nx_n^2.$$

Rodovna funkcija nenegativne celoštevilske slučajne spremenljivke X je

$$G_X(t) = \sum_n p_nt^n,$$

kjer je $p_n = P[X = n]$ za $n = 0, 1, 2, ...$Pri tem velja

$$E(X) = G'_X(1) \quad \text{in} \quad D(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2 .$$

Iz rodovne funkcije dobimo verjetnostno z razvojem v Taylorjevo vrsto

$$G_X(t) = \sum_n p_nt^n = \sum_n \frac{G'_X(0)}{n!}t^n.$$

Nekatere vrste in porazdelitve:

- binomska formula $(x + y)^n = \sum_{k=1}^n \binom{n}{k} x^ky^{n-k}$,
- končna geometrijska vrsta $1 + x + x^2 + ... + x^n = \frac{1-x^{n+1}}{1-x}$,
- geometrijska vrsta $1+x+x^2+...+x^n+... = \frac{1}{1-x}$ in njen odvod $1+2x+3x^2+...+nx^{n-1}+... = \frac{1}{(1-x)^2}$,
- binomska porazdelitev $b(n, p)$ s parametroma $n \in \mathbb{N}$ in $p \in (0, 1)$:

$$P[X = k] = \binom{n}{k} p^kq^{n-k}, \quad q = 1 - p, \quad k = 0, 1, ..., n,$$
$$E(X) = np, \quad D(X) = npq, \quad G_X(t) = (pt + q)^n ,$$

- geometrijska porazdelitev s parametrom $p \in (0, 1)$:

$$P[X = k] = pq^{k-1}, \quad q = 1 - p, \quad k = 1, 2, ..., n, ...$$
$$E(X) = \frac{1}{p}, \quad D(X) = \frac{q}{p^2}, \quad G_X(t) = \frac{pt}{1 - qt}.$$

Formule za 2. kolokvij iz verjetnosti

Slučajna spremenljivka X je **diskretna**, če je njena zaloga števna množica $\{x_1, x_2, x_3, ...\}$. Porazdelitev slučajne spremenljivke X opišemo

- z verjetnostno funkcijo:

$$X \sim \begin{pmatrix} x_1 & x_2 & x_3 & \cdots \\ p_1 & p_2 & p_3 & \cdots \end{pmatrix}, \quad p_1 + p_2 + p_3 + \cdots = 1,$$

  kjer je $P[X = x_n] = p_n$ za $n = 1, 2, 3, ...$

- s porazdelitveno funkcijo:

$$F_X(x) = P[X < x] = \sum_{x_i < x} p_i,$$

  kjer velja $\lim_{x\to\infty} F_X(x) = 1$, $\lim_{x\to-\infty} F_X(x) = 0$, in $P[X = x_0] = \lim_{x\downarrow x_0} F_X(x) - F_X(x_0)$.

Slučajna spremenljivka X je **zvezna**, če je njena porazdelitvena funkcija $F_X(x) = P[X < x]$ zvezna funkcija. Porazdelitev slučajne spremenljivke X opišemo z gostoto verjetnosti p_X, to je nenegativna realna funkcija, za katero velja

$$\int_{\mathbb{R}} p(x)\, dx = 1.$$

Pri tem velja

$$F_X(x) = P[X < x] = \int_{-\infty}^{x} p(t)\, dt \quad \text{in} \quad F_X'(x) = p_X(x),$$

$$P[a < X < b] = \int_a^b p(x)\, dx = F_X(b) - F_X(a).$$

Začetni moment reda n slučajne spremenljivke X je $z_n = E(X^n)$ in ga izračunamo v diskretnem oziroma zveznem primeru kot

$$z_n = \sum_k p_k x_k^n \quad \text{oziroma} \quad z_n = \int_{\mathbb{R}} x^n p(x)\, dx.$$

Pri tem je $z_1 = E(X)$ **matematično upanje** slučajne spremenljivke X.
Centralni moment reda n slučajne spremenljivke X je

$$m_n = E\left((X - E(X))^n\right).$$

Pri tem je $m_2 = D(X) = E\left((X - E(X))^2\right)$ **disperzija** slučajne spremenljivke X in se izračuna kot

$$D(X) = E\left(X^2\right) - E(X)^2 = z_2 - z_1^2.$$

Asimetrija slučajne spremenljivke X je definirana z

$$A(X) = \frac{m_3}{\sqrt{m_2^3}},$$

kjer upoštevamo zvezo $m_3 = z_3 - 3z_2z_1 + 2z_1^3$.
Kvartil reda i (za $i = 1, 2, 3$) q_i je pri zvezni slučajni spremenljivki X definiran s predpisom

$$F_X(q_i) = \frac{i}{4}.$$

Pri tem je $m = q_2$ **mediana** in $s = (q_3 - q_1)/2$ **semiinterkvartilni razmik**.

Zvezna slučajna spremenljivka X je porazdeljena **normalno** $\mathcal{N}(a, \sigma)$ z matematičnim upanje a in standardnim odklonom σ, če je njena gostota

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Standardizirana normalna porazdelitev $\mathcal{N}(0, 1)$ ima gostoto

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

in porazdelitveno funkcijo

$$F_X(x) = \frac{1}{2} + \Phi(x), \quad \text{kjer je} \quad \Phi(x) = \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

Za aproksimacijo binomske porazdelitve $X \sim b(n, p)$ uporabimo:

- Poissonov obrazec oziroma Laplaceovo lokalno formulo:

$$P[X = k] = \binom{n}{k} p^k q^{n-k} \approx \frac{(np)^k}{k!} e^{-np} \quad \text{oz.} \quad P[X = k] \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}}.$$

- Laplaceovo integralsko formulo

$$P[a < X < b] = P\left[\frac{a - np}{\sqrt{npq}} < X^* < \frac{b - np}{\sqrt{npq}}\right] = \Phi\left(\frac{b - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - np}{\sqrt{npq}}\right).$$

Diskretni slučajni vektor (X, Y) ima diskretno zalogo vrednosti $\{(x_i, y_j) | i = 1, 2, ..n; j = 1, 2, ..., m\}$ v $\mathbb{R}^2$. Njegovo porazdelitev praviloma podamo z verjetnostno tabelo

$X \backslash Y$	y_1	y_2	$\cdots$	y_m	
x_1	p_{11}	p_{12}	$\cdots$	p_{1m}	$p_{1.}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2m}	$p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\cdots$	p_{nm}	$p_{n.}$
	$q_{.1}$	$q_{.2}$	$\cdots$	$q_{.m}$	

,

pri tem je

$$p_{ij} = P[(X, Y) = (x_i, y_j)], \quad p_{i.} = \sum_{j=1}^{m} p_{ij}, \quad q_{.j} = \sum_{i=1}^{n} p_{ij}, \quad \sum_{j=1}^{m} \sum_{1=1}^{n} p_{ij} = \sum_{i=1}^{n} p_{i.} = \sum_{j=1}^{m} q_{.j} = 1.$$

- Robna porazdelitev slučajnih spremenljivk X in Y je

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} \quad \text{in} \quad Y \sim \begin{pmatrix} y_1 & y_2 & \cdots & y_m \\ q_1 & q_2 & \cdots & q_m \end{pmatrix}$$

- Slučajni spremenljivki X in Y sta neodvisni če za vse i in j velja

$$p_{ij} = P[(X, Y) = (x_i, y_j)] = P[X = x_i] P[Y = y_j] = p_i q_j$$

Zvezni slučajni vektor (X, Y) je podan z gostoto $p : \mathbb{R}^2 \to \mathbb{R}$, to je nenegativno realno funkcijo dveh spremenljivk, za katero velja

$$\iint_{\mathbb{R}^2} p(x, y) \, dxdy = 1.$$

Pri tem velja

$$P[(X, Y) \in D] = \iint_D p(x, y) \, dxdy.$$

- Robna porazdelitev slučajnih spremenljivk X in Y je

$$p_X(x) = \int_{\mathbb{R}} p(x, y) \, dy \quad \text{in} \quad p_Y(y) = \int_{\mathbb{R}} p(x, y) \, dx.$$

- Slučajni spremenljivki X in Y sta neodvisni če velja $p(x, y) = p_X(x) p_Y(y)$.