1	2	3	Σ

FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

OSNOVE VERJETNOSTI IN STATISTIKE 20010/2011

TEORIJA (20. SEPTEMBER 2011)

IME IN PRIIMEK: ______________________ VPISNA ŠT: | | | | | | | | |

NAVODILA

Pazljivo preberite besedila vprašanj, predno pričnete pisati odgovore. Čas pisanja je 30 minut. Možnih točk je 30, za pozitivno oceno je potrebno zbrati vsaj polovico (najmanj po 3 pri vsaki nalogi). Veliko uspeha!

1. (a) Definiraj produkt dveh dogodkov, pogojno verjetnost in (d) podaj formulo za njen izračun. (b) Kako izračunamo pogojno verjetnost dveh neodvisnih dogodkov? (c) Definiraj popoln sistem dogodkov ter (e) podaj formulo za popolno verjetnost. (f) Kako lahko izračunamo verjetnost, da je dogodek A nastopil skupaj z določenim dogodkom (hipotezo) H_i iz prve faze? (g) Za bonus izpelji še Bayesov obrazec.

[1] (a) B dogodek, $P(B)>0$: pogojna verjetnost $P(A/B)$ je verjetnost dogodka A, pri čemer smo prejšnjemu kompleksu pogojev K pridružili še, da se je zgodil dogodek B: $K' = K \cap B$.

[1] (d) $P(A/B) = \dfrac{P(AB)}{P(B)}$ (*)   [1] $AB = A\cap B$ (produkt dogodkov A in B) je dogodek, ki se zgodi, če se zgodita A in B hk

[1] (c) $\{A_1, \dots, A_n\}$ je popoln sistem dogodkov: tj. množica nepraznih dogodkov, pri katerih se v vsaki ponovitvi poskusa zgodi natanko eden med njimi.

$A_i \neq N \quad i = 1, \dots, n$

$A_i \cap A_j = N \quad \text{za } i \neq j$

$A_1 \cup \dots \cup A_n = G$

(b) [2] Za neodvisna dogodka A in B je $P(A\cap B) = P(A)P(B)$ in zato $P(A/B) = P(A)$

Za popoln sistem dogodkov $\{H_1, \dots, H_n\}$ in poljuben dogodek A velja

[1] (e)
$$P(A) = \sum_{i=1}^{n} P(A\cap H_i) = \sum_{i=1}^{n} P(H_i)P(A/H_i). \quad (*)$$

[2] (f)
$$P(H_k/A) = \frac{P(H_k)P(A/H_k)}{\sum_{i=1}^{n} P(H_i)P(A/H_i)}$$
Bayesov obrazec

[3] (g) Izpeljava

Iz (*): $P(AB) = P(B)P(A/B)$

in simetrično $= P(A)P(B/A)$

$\Rightarrow$ $P(A)P(B/A) = P(B)P(A/B)$ oziroma za $B = H_k$ dobimo

$P(A)P(H_k/A) = P(H_k)P(A/H_k)$

in končno $P(H_k/A) = \dfrac{P(H_k)P(A/H_k)}{P(A)}$

preostane še uporaba (*)

a b c

2. Definiraj matematično upanje in standardni odklon slučajne spremenljivke. Opiši postopek za standardizacijo slučajne spremenljivke in zapiši njeno matematično upanje ter njen odklon. Naj bo X slučajna spremenljivka z $\mathsf{E}X = \mu$ in $\mathsf{D}X = \sigma^2$. (d) Za njen slučajen vzorec $\{X_i\}_{i=1}^{n}$ definiraj vzorčno povprečje $\overline{X}$ in (e) napiši, kaj se dogaja z vzorčnim povprečjem $\mathsf{E}\bar{X}$ in (f) s standardno napako $\mathsf{D}\bar{X}$ z naraščanjem velikosti vzorca.

(a)
$$\mu = EX = \begin{cases} \sum_{i=1}^{\infty} x_i p_i, & \text{če je } X \text{ diskretna sl. spr.} \\ \int_{-\infty}^{\infty} x\,p(x)\,dx, & \text{če je } X \text{ zv. sl. sp.} \end{cases}$$ [4]

pri čemer je $p_i = P(X = x_i)$ in $p(x)$ gostota verjetnosti sl. spr. X.

EX obstaja, če $\infty > \sum_{i=1}^{\infty} |x_i| p_i$ / $\int_{-\infty}^{\infty} |x|\,p(x)\,dx$ [1]

(b) [1]
$$\sigma^2 = DX = E(X^2) - (EX)^2$$ in obstaja, če obstaja $E(X^2)$.

(c)
Std: $Z = \dfrac{X-\mu}{\sigma}$ [3], $EZ = 0$, $DZ = 1$

Vzorec $(X_1, \dots, X_n)$ in $EX_i = \mu$, $DX_i = \sigma^2$, X_i med seboj neodvisni

[1] (d) $\bar{X} = \dfrac{X_1 + \dots + X_n}{n}$ $\Rightarrow$ (e) $E\bar{X} = \dfrac{EX_1 + \dots + EX_n}{n} = \dfrac{n\mu}{n} = \mu \xrightarrow[n\to\infty]{} \mu$

linearnost mat. upanja

[3]

(f) $D\bar{X} = \dfrac{DX_1 + \dots + DX_n}{n^2} = \dfrac{n\sigma^2}{n^2} = \dfrac{\sigma^2}{n} \xrightarrow[n\to\infty]{} 0$

X_i neodvisni $\Rightarrow$ nekorelirani

a b

3. Pojasni razliko med točkovno in intervalno oceno? Opiši postopek intervalskega ocenjevanja parametrov in (c) pojasni, kaj nam pove koeficienta zaupanja $(1-\alpha)$ (teoretična interpretacija).
(d) Opiši ocenjevanje parametrov z majhnimi vzorci (čim več možnosti).

Točkovna cenilka je formula (pravilo), ki nam pove kako izračunati num. oceno
[3] (a) parametra populacije na osnovi vzorca in rezultatu ne moremo zaupati v smislu ve
Pri intervalni oceni pa znamo oceniti verjetnost, da parameter populacije je na izračunanem intervalu ($P(a < \mu < b) = 1-\alpha$)

(b) S slučajnim vzorcem ocenjujemo parameter μ populacije. Poskušamo najti
(1) statistiko g (formula/cenilko), ki je nepristranska (tj. $Eg = \mu$) in se na vseh možnih vzorcih vsaj približno normalno porazdeljuje s standardno napako $SE(g)$, in
[4] (2) t.i. interval, v katerem se bo z dano gotovostjo $(1-\alpha)$ nahajal ocenjevani paramete
(a in b sta spodnja in zg. meja zaupanja, α pa stopnja tveganja)
(3) Izberemo ustrezni test.
(4) Za vsak enostavni slučajni vzorec lahko izračunamo ob izbrani stopnji tveganja α interval zaupanja za parameter μ (meji sta slučajni spremenljivki).

(c) (Teoretična) interpretacija intervala zaupanja: z verjetnostjo tveganja α se parameter μ nahaja v tem intervalu. Če zaporedoma izbiramo vzorce velikosti n in za vsakega izračunamo interval zaupanja za parameter μ, tedaj pričakujemo, da bo $(1-\alpha)\cdot 100\%$ intervalov vsebovalo pravo vrednost parametra.

(d) Majhni vzorci: (1) ali se sl. spremenljivka porazdeljuje normalno (DA/N
(2) ali poznamo standardni odklon σ (DA/NE
(3) kakšna je velikost vzorca [3]

DA/DA z
NE/DA t
DA/NE velik vzorec ($n > 30$) χ^2
DA/NE mali vzorec ($n < 30$)