STATISTIKA 11. 5. 2012, drugi test, teorija. Odgovori in pojasnila

1. (3) Navedite **definicijo intervala zaupanja** za neznan parameter p porazdelitve statističnega znaka pri stopnji zaupanja $1-\alpha$. Potem še s svojimi besedami opišite postopek **ocenjevanja** z intervalom zaupanja. Kakšne so prednosti glede na ocenjevanje s cenilkami?

[Odg.] Definicija: glejte predavanja oz. povzetke.

Za ocenjevanje z intervalom zaupanja potrebujemo podatke, formule (ali rač. program) za izračun statistik L in D, pri tem moramo morda poznati tudi njuni porazdelitvi in uporabiti ustrezne tabele "kritičnih" vrednosti).

Pri velikih stopnjah zaupanja obstaja velika verjetnost, da bomo izračunali interval, ki vsebuje pravo vrednost parametra. Dobimo tudi informacijo o "razpršenosti" možnih vrednosti parametra.

2. (3) Kakšno statistično domnevo preverjamo s **prilagoditvenimi** testi? Opišite korake testiranja s testom χ^2. Če vemo, da je porazdelitev znaka U **normalna**, naštejte teste, s katerimi bi lahko testirali domnevo, da je U porazdeljena po zakonu $N(100,\ 10)$.

(Prosim, da svoj odgovor napišete na **hrbtni** strani tega lista.)

[Odg.] S prilag. testi preverjamo domnevo o tipu (vrsti) porazdelitvenega zakona statistične spremenljivke. Prilagoditveni test χ^2 : glejte predavanja oz. povzetke. Testiranje $U \sim N(100,\ 10)$: preverjamo lahko s testom Kolmogorova ali s prilagoditvenim testom χ^2, lahko pa tudi s parom parametričnih testov: test o $E(U) = 100$ in test o $\sigma = 10$.

Obkrožite črke pred pravilnimi izjavami. Pravilna odločitev: 2, nepravilna: -1, brez nje: 0 točk.

3. V zaporedju neodvisnih naključnih spr. K_n so vse porazdeljene po zakonu $b(72,\ 1/3)$.

a K_{10} lahko dobro aproksimiramo s Poissonovo porazdelitvijo $P(24)$.

b K_{10} lahko dobro aproksimiramo s porazdelitvijo $N(24,\ 16)$.

c Limitna porazdelitev zaporedja K_n je normalna $N(24,\ 16)$.

d Limitna porazdelitev zaporedja $X_m = (K_1 + K_2 + \ldots + K_m)/m$ je konstanta 24.

e Limit. porazd. zap. $Y_m = (K_1 + K_2 + \ldots + K_m - 24m)/\sqrt{m}$ je normalna $N(0,1)$.

[Pojasnila]

Za vsako K_n velja: $E(K_n) = 72.1/3 = 24$ in $D(K_n) = 72.1/3.2/3 = 16$, zato $\sigma(K_n) = 4$.

Pogoj za dobro aproksimacijo porazdelitve $b(n,p)$ s Poissonovo porazdelitvijo $P(\lambda)$ je: p blizu 0 (in q blizu 1) oz. $\lambda = np \approx npq$ (za $b(n,p)$). Ta pogoj ni izpolnjen, 24 ni blizu 16, zato a ne velja.

Binomsko porazdelitev $b(n,p)$ lahko aproksimiramo z normalno porazdelitvijo $N(np,\ \sqrt{npq})$ pri velikih n, zaradi $\sqrt{npq} = 4$ torej b ne velja.

Vse porazdelitve K_n so enake, torej je tudi limitna porazdelitev takšna, torej binomska $b(72,\ 1/3)$ in c tudi ne velja.

Zaradi neodvisnosti je $S_m = mX_m = K_1 + \ldots + K_m$ spet porazdeljena binomsko, in sicer $b(m.72,\ 1/3)$. Ko je m velik, lahko to porazdelitev aproksimiramo z normalno $N(m.24,\ 4\sqrt{m})$. Potem je $X_m \sim N(24,\ 4/\sqrt{m})$ in zaradi $4/\sqrt{m} \longrightarrow 0$ velja d.

Spremenljivke S_m so porazdeljene normalno, če jih standardiziramo, dobimo

$Z_m = (S_m - E(S_m))/\sigma(S_m) = (K_1 + \ldots + K_m - 24.m)/4\sqrt{m}$,

zato zaporedje Y_m konvergira proti $N(0,4)$ in e ne velja.

4. X in Y sta naključni spremenljivki z omejenima zalogama vrednosti, ki nastopata v istem poskusu.

a Če sta neodvisni, sta tudi nekorelirani.

b Če sta nekorelirani, sta tudi neodvisni.

c Iz njune neodvisnosti sledi, da je $r(X, Y) = 0$.

d Če sta enako porazdeljeni, je $r(X, Y) = 1$.

e Iz $r(X, Y) = -1$ sledi, da je $Y = -X$.

[Pojasnila]

Spremenljivki sta omejeni, zato obstajajo vsi njuni momenti, torej imata (končni) matematični upanji in disperziji, pa tudi korelacijski koeficient obstaja.

Iz neodvisnosti X in Y v takšnem primeru sledi $E(XY) = E(X)E(Y)$ in $K(X, Y) = 0$, zato a velja.

Iz primerov (tudi na predavanjih) je znano, da b v splošnem ne velja.

Pravilnost a ima za posledico pravilnost b.

Tudi enako porazdeljeni spremenljivki sta lahko neodvisni, torej tudi nekorelirani, zato d ne velja.

Iz $r(X, Y) = -1$ sledi le, da je $Y = aX + b$, kjer je število a negativno. Torej e ne velja.