TESTIRANJE STATISTIČKIH HIPOTEZA

Statistička hipoteza je pretpostavka koja se odnosi na raspodelu obeležja X ili na neki parametar u raspodeli obeležja koja se proverava na uzorku. Sa H_0 obeležavamo nultu hipotezu koja se proverava pomoću realizovanog uzorka. Postupak prihvatanja ili odbacivanja hipoteze naziva se testiranjem i vrši se po unapred utvrđenom pravilu koje za svaki dobijeni uzorak određuje kako da se postupi (odbaciti ili ne odbaciti H_0). Pravilo za testiranje hipoteze naziva se statističkim testom. Statistika na osnovu koje se testiranje izvodi naziva se test statistikom. Hipoteza H_1 koju suprostavljamo nultoj hipotezi zove se suprotnom ili alternativnom hipotezom.

Pre uzimanja uzorka i testiranja, utvrđuje se prag značajnosti α koji predstavlja verovatnoću sa kojom se hipoteza H_0 odbacuje ako je tačna. Na osnovu praga značajnosti i raspodele test statistike određuje se kritična oblast *C*. Kritična oblast se najčešće iskazuje preko kritičnih vrednosti test statistike. Hipoteza H_0 se odbacuje ako vrednost koju test statistika primi u realizovanom uzorku pripada *C*; u suprotnom, hipoteza H_0 se ne odbacuje.

Statistički testovi kod kojih raspodela test statistike bitno zavisi od raspodele posmatranog obeležja *X* nazivaćemo parametarskim testovima, a testovi kod kojih raspodela test statistike ne zavisi od raspodele obeležja *X* nazivaćemo neparametarskim testovima.

PARAMETARSKI TESTOVI ZA NORMALNU RASPODELU

Testiranje hipoteze $H_0(a = a_0)$ o matematičkom očekivanju a kada je disperzija σ^2 poznata

Test statistika je $U = \frac{(\bar{X}-a_0)\sqrt{n}}{\sigma}$. Ako je hipoteza $H_0(a = a_0)$ tačna, onda test statistika ima $N(0,1)$ raspodelu ako obeležje X ima $N(a, \sigma^2)$ raspodelu ili približno $N(0,1)$ ako je uzorak veliki, a obeležje *X* nema normalnu raspodelu.

Za dati prag značajnosti α određuje se kritična oblast C_α u zavisnosti od alternativne hipoteze.

1. Ako je $\boldsymbol{H_1: a \neq a_0}$ iz uslova $P\left\{\left|\frac{(\bar{X}-a_0)\sqrt{n}}{\sigma}\right| \geq c\right\} = \alpha = 1 - 2\Phi(c)$, konstantu c nalazimo u tablici IV za normalnu raspodelu iz relacije $\Phi(c) = \frac{1-\alpha}{2}$. Kritična oblast je $C_\alpha = (-\infty, -c] \cup [c, +\infty)$.
2. Ako je $\boldsymbol{H_1: a > a_0}$ iz uslova $P\left\{\frac{(\bar{X}-a_0)\sqrt{n}}{\sigma} \geq c_1\right\} = \alpha = 0.5 - \Phi(c_1)$, konstantu c_1 nalazimo u tablici IV za normalnu raspodelu. Kritična oblast je $C_\alpha = [c_1, +\infty)$.
3. Ako je $\boldsymbol{H_1: a < a_0}$ iz uslova $P\left\{\frac{(\bar{X}-a_0)\sqrt{n}}{\sigma} \leq -c_2\right\} = \alpha = 0.5 - \Phi(c_2)$, konstantu c_2 nalazimo u tablici IV za normalnu raspodelu. Kritična oblast je $C_\alpha = (-\infty, -c_2]$.

Na realizovanom uzorku izračunavamo numeričku vrednost statistike $u = \frac{\bar{x}-a_0}{\sigma}\sqrt{n}$. Ako $u \in C_\alpha$, tada hipotezu H_0 odbacujemo, ako $u \notin C_\alpha$ tada hipotezu H_0 ne odbacujemo.

Testiranje hipoteze $H_0(a = a_0)$ o matematičkom očekivanju a kada disperzija σ^2 nije poznata

Test statistika je $U = \frac{(\bar{X}-a_0)\sqrt{n-1}}{S}$ ili $U = \frac{(\bar{X}-a_0)\sqrt{n}}{\tilde{S}}$. Ako je hipoteza $H_0(a = a_0)$ tačna i ako pretpostavimo da obeležje X ima $N(a, \sigma^2)$ raspodelu, test statistika ima Studentovu raspodelu sa $n-1$ stepenom slobode (tablica VI - $P\{|t_n| > t\} = \alpha$).

Za dati prag značajnosti α određuje se kritična oblast C_α u zavisnosti od alternativne hipoteze.

1. Ako je $\boldsymbol{H_1: a \neq a_0}$ iz uslova $P\left\{\left|\frac{(\bar{X}-a_0)}{S}\sqrt{n-1}\right| \geq t_\alpha\right\} = \alpha$, konstantu t_α nalazimo u tablici za Studentovu raspodelu sa $n-1$ stepenom slobode. Kritična oblast je $C_\alpha = (-\infty, -t_\alpha] \cup [t_\alpha, +\infty)$.
2. Ako je $\boldsymbol{H_1: a > a_0}$ iz uslova $P\left\{\frac{(\bar{X}-a_0)}{S}\sqrt{n-1} \geq t_{2\alpha}\right\} = \alpha$, konstantu $t_{2\alpha}$ nalazimo u tablici za Studentovu raspodelu sa $n-1$ stepenom slobode. Kritična oblast je $C_\alpha = [t_{2\alpha}, +\infty)$.
3. Ako je $\boldsymbol{H_1: a < a_0}$ iz uslova $P\left\{\frac{(\bar{X}-a_0)}{S}\sqrt{n-1} \leq -t_{2\alpha}\right\} = \alpha$, konstantu $t_{2\alpha}$ nalazimo u tablici za Studentovu raspodelu sa $n-1$ stepenom slobode. Kritična oblast je $C_\alpha = (-\infty, -t_{2\alpha}]$.

Na realizovanom uzorku izračunavamo numeričku vrednost statistike $\mathrm{u} = \frac{\bar{\mathrm{x}}-a_0}{\mathrm{s}}\sqrt{\mathrm{n}-1}$. Ako $u \in C_\alpha$, tada hipotezu H_0 odbacujemo, ako $u \notin C_\alpha$ tada hipotezu H_0 ne odbacujemo.

Testiranje hipoteze $H_0(\sigma^2 = {\sigma_0}^2)$ o disperziji σ^2 ako matematičko očekivanje a nije poznato

Iz pretpostavke da obeležje X ima normalnu raspodelu $N(a, \sigma^2)$ sledi da test statistika $U = \frac{nS^2}{{\sigma_0}^2}\left(= \frac{(n-1)\tilde{S}^2}{{\sigma_0}^2}\right)$ ima χ^2 raspodelu sa $n-1$ stepenom slobode ako je hipoteza H_0 tačna.

Za dati prag značajnosti α određuje se kritična oblast C_α u zavisnosti od alternativne hipoteze.

1. Ako je $\boldsymbol{H_1: \sigma^2 \neq {\sigma_0}^2}$ iz uslova $P\{U \leq c_1\} = \frac{\alpha}{2}$ i $P\{U \geq c_2\} = \frac{\alpha}{2}$, nalazimo u tablici za χ^2 raspodelu sa $n-1$ stepenom slobode (tablici V) vrednosti za c_1 i c_2. Kritična oblast je $C_\alpha = [0, c_1] \cup [c_2, +\infty)$.
2. Ako je $\boldsymbol{H_1: \sigma^2 > {\sigma_0}^2}$ iz uslova $P\{U \geq c\} = \alpha$, nalazimo u tablici V vrednost za c. Kritična oblast je $C_\alpha = [c, +\infty)$.
3. Ako je $\boldsymbol{H_1: \sigma^2 < {\sigma_0}^2}$ iz uslova $P\{U \leq c\} = 1 - \alpha$, nalazimo u tablici V vrednost za c. Kritična oblast je $C_\alpha = (0, c)$.

Testiranje hipoteze $H_0(\sigma^2 = {\sigma_0}^2)$ o disperziji σ^2 ako je matematičko ocekivanje a poznato

Za test statistiku koristimo $U = \frac{n\bar{s}^2}{{\sigma_0}^2}$. Ako je hipoteza H_0 tačna, tada iz pretpostavke da obeležje X ima normalnu raspodelu $N(a, \sigma^2)$, sledi da statistika U ima χ_n^2 raspodelu. Postupak testiranja sprovodimo kao u slučaju kada matematičko očekivanje nije poznato, samo što se sada koristi χ_n^2 raspodela.

Za **veliki uzorak** $(n > 30)$ koristimo aproksimaciju χ_n^2 (χ_{n-1}^2) raspodele normalnom $N(n, 2n)$ ($N\big(n-1, 2(n-1)\big)$) raspodelom. Tada $(\chi_n^2)^* = \frac{\chi_n^2 - n}{\sqrt{2n}} \sim N(0,1)$, pa se c određuje iz tablice za normalnu $N(0,1)$ raspodelu.