Formule za 6. vajo pri predmetu Naključni pojavi

Statistična hipoteza je trditev o parametru ali verjetnostni porazdelitvi ene ali več populacij. Če se hipoteza nanaša na parameter, jo imenujemo parametrična, če pa se nanaša na porazdelitev, jo imenujemo neparametrična. Preverjanje hipoteze je postopek ugotavljanja njene pravilnosti. Postopek imenujemo statistični test. V tej vaji obravnavamo preverjanje parametričnih hipotez.
Preverjano hipotezo imenujemo ničelna in označimo H_0, njej nasprotujočo hipotezo pa imenujemo alternativna in označimo H_1. Ničelna hipoteza vedno trdi, da je parameter (npr. povprečje populacije) enak neki vrednosti, npr.:

$$\mathrm{H}_0 : m = 20\,\mathrm{g}\,, \tag{1}$$

alternativna pa, da je bodisi neenak (dvostranska), manjši (leva enostranska) bodisi večji (desna enostranska) od te vrednosti, npr.:

$$\mathrm{H}_1 : m \neq 20\,\mathrm{g}\,, \qquad \text{ali} \qquad \mathrm{H}_1 : m < 20\,\mathrm{g}\,, \qquad \text{ali} \qquad \mathrm{H}_1 : m > 20\,\mathrm{g}\,. \tag{2}$$

Hipoteze preverjamo v osmih korakih:

1. Na podlagi problema izberemo parameter porazdelitve naključne spremenljivke, katerega vrednost preverjamo.
2. Za izbrani parameter postavimo ničelno hipotezo H_0.
3. Glede na problem postavimo alternativno hipotezo H_1, ki je lahko dvostranska ali leva ali desna enostranska.
4. Izberemo stopnjo značilnosti testa α, običajno $\alpha = 0.05$ ali 0.01.
5. Na podlagi cenilke parametra, ki jo uporabljamo, izberemo primerno testno statistiko.
6. Za izbrano testno statistiko določimo interval zavračanja oz. sprejemanja ničelne hipoteze glede na izbrano alternativno hipotezo. Meji intervala imenujemo kritični vrednosti testne statistike.
7. Izračunamo vrednost testne statistike.
8. Glede na vrednost testne statistike se odločimo o veljavnosti ničelne hipoteze in sklep utemeljimo.

Pri testiranju hipotez so glede na dejansko veljavnost H_0 možne naslednje štiri situacije:

	H_0 drži	H_0 ne drži
H_0 zavrnemo	napaka 1. vrste	pravilna odločitev
H_0 ne zavrnemo	pravilna odločitev	napaka 2. vrste

Napako 1. vrste imenujemo tudi tveganje proizvajalca (npr. da je ustrezna serija pri statističnem testu spoznana kot neustrezna in zato zavrnjena) in jo običajno izberemo pred izvajanjem testa. Verjetnost zanjo je:

$$P\,(\text{napaka 1. vrste}) = \alpha\,. \tag{3}$$

Napako 2. vrste imenujemo tudi tveganje kupca (npr. da je sprejeta neustrezna serija, ki pa je bila pri testu spoznana kot ustrezna) in je ne moremo določiti vnaprej, ker je odvisna od dejanskega stanja preverjane populacije. Verjetnost zanjo je:

$$P\,(\text{napaka 2. vrste}) = \beta\,. \tag{4}$$

Testiranje hipoteze za povprečje normalno porazdeljene populacije z znano varianco (dovoljen poljuben n) ali poljubno porazdeljene populacije z neznano varianco (potreben $n > 30$, ocenimo $\sigma^2 = S^2$):

H_0	H_1	testna statistika	območje zavračanja H_0
$m = m_0$	$m < m_0$ $m \neq m_0$ $m > m_0$	$Z = (\langle X \rangle - m_0) \,/\, (\sigma/\sqrt{n})$	$z < -z_\alpha$ $z < -z_{\alpha/2} \cup z > z_{\alpha/2}$ $z > z_\alpha$

Testiranje hipoteze za povprečje normalno porazdeljene populacije z neznano varianco (na voljo $n < 30$):

H_0	H_1	testna statistika	območje zavračanja H_0
$m = m_0$	$m < m_0$ $m \neq m_0$ $m > m_0$	$T = (\langle X \rangle - m_0) \,/\, (S/\sqrt{n})$	$t < -t_{n-1;\alpha}$ $t < -t_{n-1;\alpha/2} \cup t > t_{n-1;\alpha/2}$ $t > t_{n-1;\alpha}$

Testiranje hipoteze za varianco normalno porazdeljene populacije:

H_0	H_1	testna statistika	območje zavračanja H_0
$\sigma^2 = \sigma_0^2$	$\sigma^2 < \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$ $\sigma^2 > \sigma_0^2$	$\chi^2 = (n-1)S^2/\sigma_0^2$	$\chi^2 < \chi^2_{n-1;1-\alpha}$ $\chi^2 < \chi^2_{n-1;1-\alpha/2} \cup \chi^2 > \chi^2_{n-1;\alpha/2}$ $\chi^2 > \chi^2_{n-1;\alpha}$

Testiranje hipoteze za delež populacije, če lahko binomsko porazdelitev aproksimiramo z normalno:

H_0	H_1	testna statistika	območje zavračanja H_0
$p = p_0$	$p < p_0$ $p \neq p_0$ $p > p_0$	$Z = (p - p_0) \,/\, \sqrt{p_0(1-p_0)/n}$	$z < -z_\alpha$ $z < -z_{\alpha/2} \cup z > z_{\alpha/2}$ $z > z_\alpha$

Testiranje hipoteze za vsoto (razliko) povprečij normalno porazdeljenih populacij z znanima variancama (dovoljena poljubna n_1 in n_2) ali poljubno porazdeljenih populacij z neznanima variancama (potrebna $n_1 > 30$ in $n_2 > 30$, ocenimo $\sigma_1^2 = S_1^2$, $\sigma_2^2 = S_2^2$):

H_0	H_1	testna statistika	območje zavračanja H_0
$m_1 \pm m_2 = \Delta_0$	$m_1 \pm m_2 < \Delta_0$ $m_1 \pm m_2 \neq \Delta_0$ $m_1 \pm m_2 > \Delta_0$	$Z = [(\langle X_1 \rangle \pm \langle X_2 \rangle) - \Delta_0] \,/\, \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$	$z < -z_\alpha$ $z < -z_{\alpha/2} \cup z > z_{\alpha/2}$ $z > z_\alpha$

Testiranje hipoteze za vsoto (razliko) povprečij normalno porazdeljenih populacij z neznanima, a podobnima variancama (na voljo $n_1 < 30$ in $n_2 < 30$):

H_0	H_1	testna statistika	območje zavračanja H_0
$m_1 \pm m_2 = \Delta_0$	$m_1 \pm m_2 < \Delta_0$ $m_1 \pm m_2 \neq \Delta_0$ $m_1 \pm m_2 > \Delta_0$	$T = [(\langle X_1 \rangle \pm \langle X_2 \rangle) - \Delta_0] \,/\, \left(S_p\sqrt{1/n_1 + 1/n_2}\right)$ $S_p^2 = \left((n_1-1)S_1^2 + (n_2-1)S_2^2\right) / (n_1+n_2-2)$	$t < -t_{n_1+n_2-2;\alpha}$ $t < -t_{n_1+n_2-2;\alpha/2} \cup t > t_{n_1+n_2-2;\alpha/2}$ $t > t_{n_1+n_2-2;\alpha}$

Testiranje hipoteze za vsoto (razliko) deležev populacij, pri katerih lahko binomski porazdelitvi aproksimiramo z normalnima:

H_0	H_1	testna statistika	območje zavračanja H_0
$p_1 \pm p_2 = \Delta_0$	$p_1 \pm p_2 < \Delta_0$ $p_1 \pm p_2 \neq \Delta_0$ $p_1 \pm p_2 > \Delta_0$	$Z = ((p_1 \pm p_2) - \Delta_0) \,/\, \sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}$	$z < -z_\alpha$ $z < -z_{\alpha/2} \cup z > z_{\alpha/2}$ $z > z_\alpha$

Testiranje hipoteze za razmerje varianc normalno porazdeljenih populacij:

H_0	H_1	testna statistika	območje zavračanja H_0
$\sigma_1^2/\sigma_2^2 = \Delta_0$	$\sigma_1^2/\sigma_2^2 < \Delta_0$ $\sigma_1^2/\sigma_2^2 \neq \Delta_0$ $\sigma_1^2/\sigma_2^2 > \Delta_0$	$F = (S_1^2/\sigma_1^2) \,/\, (S_2^2/\sigma_2^2)$	$f < 1/f_{n_2-1,n_1-1;\alpha}$ $f < 1/f_{n_2-1,n_1-1;\alpha/2} \cup f > f_{n_1-1,n_2-1;\alpha/2}$ $f > f_{n_1-1,n_2-1;\alpha}$