Preverjanje statističnih hipotez

Na osnovi danega vzorca bi radi testirali neko hipotezo o populaciji. Dani so:

- ničelna hipoteza $H_0\colon \theta = a$,
- alternativna hipoteza:
  - dvostranska: $H_1\colon \theta \neq a$,
  - zgornja enostranska: $H_1\colon \theta > a$,
  - spodnja enostranska: $H_1\colon \theta < a$,
- podatek $\theta = \theta_0$ o vzorcu velikost n,
- meje kritičnega območja:
  - dvostr. alt. hip.: θ_1 in θ_2 (ničelno hipotezo obdržimo, če $\theta_1 \leq \theta_0 \leq \theta_2$),
  - zg. enostr. alt. hipoteza: θ_2 (ničelno hipotezo obdržimo, če $\theta_0 \leq \theta_2$),
  - sp. enostr. alt. hipoteza: θ_1 (ničelno hipotezo obdržimo, če $\theta_0 \geq \theta_1$).

Pri preverjanju statističnih hipotez o populaciji sta možni dve vrsti napak:

- napaka **I. vrste** (imenujemo jo tudi **stopnja značilnosti** ali α**-napaka**) in
- napaka **II. vrste** (imenujemo jo tudi β**-napaka**).

	H_0 pravilna	H_0 napačna
ne zavrni H_0	✓	β-napaka
zavrni H_0	α-napaka	✓

Če ničelno hipotezo H_0 na osnovi danega vzorca zavrnemo, čeprav je bila H_0 pravilna (tj. za populacijo velja $\theta = a$), smo pri tem naredili napako I. vrste:

$$\alpha = P[\textbf{zavrnemo } H_0 \mid H_0 \textbf{ je pravilna}],$$
$$1 - \alpha = P[\textbf{ne zavrnemo } H_0 \mid H_0 \textbf{ je pravilna}].$$

Za slučajno spremenljivko X z vrednostjo $\theta = \theta_0$ za dani vzorec in posamezne vrste alternativnih hipotez to pomeni naslednje:

- dvostranska alternativna hipoteza: $\alpha = P[X < \theta_1 \mid \theta = a] + P[X > \theta_2 \mid \theta = a]$,
- zgornja enostranska alternativna hipoteza: $\alpha = P[X > \theta_2 \mid \theta = a]$,
- spodnja enostranska alternativna hipoteza: $\alpha = P[X < \theta_1 \mid \theta = a]$.

Izračun zgornjih verjetnosti je odvisen od opazovane lastnosti in porazdelitve slučajne spremenljivke X. Želimo si, da je α-napaka čim manjša. Napako zmanjšamo, če:

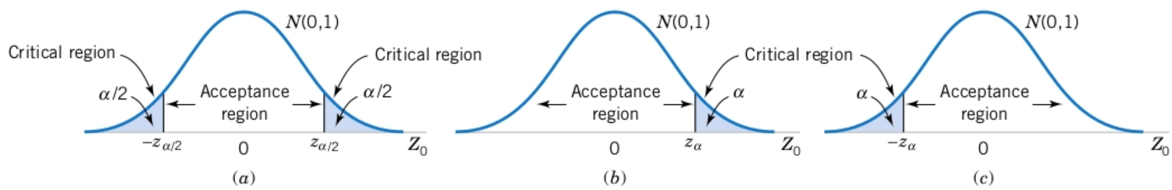

Figure 9-6 The distribution of Z_0 when H_0: $\mu = \mu_0$ is true, with critical region for (a) the two-sided alternative $H_1 : \mu \neq \mu_0$, (b) the one-sided alternative $H_1 : \mu > \mu_0$, and (c) the one-sided alternative $H_1 : \mu < \mu_0$.

- zožimo območje zavrnitve hipoteze oz. kritično območje,
- vzamemo večji vzorec.

Če ničelne hipoteze H_0 na osnovi danega vzorca **ne** zavrnemo in je bila H_0 napačna (tj. za populacijo velja $\theta \neq a$), smo pri tem naredili napako II. vrste

$$\beta = P[\textbf{ne zavrnemo } H_0 \mid \ H_0 \textbf{ je napačna}],$$
$$1-\beta = P[\textbf{zavrnemo } H_0 \mid \ H_0 \textbf{ je napačna}]$$

pa imenujemo **moč** statističnega testa.

Za slučajno spremenljivko X z vrednostjo θ_0 za dani vzorec in posamezne vrste alternativnih hipotez to pomeni naslednje:

– dvostranska alternativna hipoteza: $\beta = P[\theta_1 \leq X \leq \theta_2, \theta \neq a]$,
– zgornja enostranska altetrnativna hipoteza: $\beta = P[X \leq \theta_2, \theta \neq a]$,
– spodnja enostranska alternativna hipoteza: $\beta = P[X \geq \theta_1, \theta \neq a]$.

Pri izračunu β-napake običajno potrebujemo več informacij o populaciji, saj $\theta \neq a$. Želimo si, da je β-napaka čim manjša, moč $1-\beta$ pa čim večja. Omenimo še, da z manjšanjem α-napake raste β-napaka in obratno (če se velikost vzorca ne spreminja).

P-vrednost statističnega testa

***P*-vrednost** statističnega testa ali tudi **statistična značilnost** je najmanjša stopnja značilnosti (α-napaka), pri kateri pri danem vzorcu še zavrnemo H_0, čeprav je pravilna. Izračunamo jo kot α-napako pri najožjem možnem kritičnem območju. Za vzorčno vrednost θ_0 to pomeni:

– dvostranska alternativna hipoteza: eno mejo kritičnega območja $(-\infty, \theta_1) \cup (\theta_2, \infty)$ naj določa θ_0, druga pa naj leži simetrično glede na a,
– zgornja enostranska alternativna hipoteza: $\theta_2 = \theta_0$,
– spodnja enostranska alternativna hipoteza: $\theta_1 = \theta_0$.

Širše kritično območje bi vrednost α povečalo in ničelno hipotezo H_0 bi ob danem θ_0 zavrnili. Če pa bi bilo kritično območje ožje in ne bi vsebovalo θ_0, ničelne hipoteze H_0 ne bi zavrnili. Ničelno hipotezo torej zavrnemo le za $\alpha > P-$vrednost.

Testiranje hipotez je tesno povezano z intervali zaupanja. Na primer, če je $[l, u]$ dvostranski interval zaupanja za parameter θ (in vzorčno vrednost θ_0) s stopnjo zaupanja $1 - \alpha$, bo testiranje hipoteze

$$\begin{aligned} H_0 &\colon \theta = a, \\ H_1 &\colon \theta \neq a \end{aligned}$$

vodilo do zavrnitve H_0 natanko tedaj, ko a ne bo v $[l, u]$.

Uporabna statistika 2013/2014

VAJA 10: Preverjanje statističnih hipotez

1. V grafičnem vmesniku programskega paketa R odprite novo skriptno datoteko.
2. Tovarna, ki izdeluje samodejne termostatske regulatorje temperature, želi testirati ujemanje izdelanih regulatorjev z uradno specifikacijo izdelka. V ta namen so izbrali naključni vzorec regulatorjev in izmerili njihove preklopne temperature. Rezultati meritev so podani v datoteki 'regulatorji.txt'.

   V specifikaciji izdelkov piše, da se regulatorji preklopijo pri temperaturi $40\,^\circ C$ $(\pm 0.5\,^\circ C)$.

   a.) Kakšna je velikost n vzorca?

   b.) Izračunajte povprečno vrednost x_p in standardni odklon s preklopnih temperatur vzorca.

   c.) Narišite histogram vzorca z deležem vrednosti na ordinatni osi. Na sliko dodajte normalno krivuljo z $\mu = x_p$ in $\sigma = s$. Ali so podatki iz vzorca v skladu s pričakovanji?

   d.) Postavite ustrezni ničelno $H_0 : \mu = a$ in alternativno hipotezo H_1 o povprečni preklopni temperaturi termostatskih regulatorjev.

   e.) Določite primerno kritično območje oz. območje zavrnitve hipoteze H_0.

   f.) Predpostavimo, da so preklopne temperature regulatorjev porazdeljene normalno s standardnim odklonom $\sigma = 2$. Tedaj ima po centralnem limitnem izreku slučajna spremenljivka $Z = \frac{\bar{X} - a}{\sigma/\sqrt{n}}$ približno standardno normalno porazdelitev in je, na primer,

   $$\alpha = P(\bar{X} \notin [x_1, x_2] \mid \mu = a) = P(Z \notin [z_1, z_2]),$$

   kjer sta

   $$z_1 = \frac{x_1 - a}{\sigma/\sqrt{n}} \text{ in } z_2 = \frac{x_2 - a}{\sigma/\sqrt{n}}.$$

   Izračunajte α-napako testiranja hipoteze H_0. Ali hipotezo H_0 zavrnemo?

   g.) Izračunajte P-vrednost statističnega testa iz prejšnje točke.

h.) Določite interval zaupanja za povprečno vrednost x_p s stopnjo zaupanja $95\,\%$, če je standardni odklon populacije regulatorjev $\sigma = 2$.

i.) Predpostavimo, da standardnega odklona preklopnih temperatur populacije regulatorjev ne poznamo. Če je vzorec majhen (?), je tedaj slučajna spremenljivka $T = \frac{\bar{X}-a}{s/\sqrt{n}}$ približno t-porazdeljena z $n-1$ prostostnimi stopnjami in je, na primer,

$$\alpha = P(\bar{X} \notin [x_1, x_2] \mid \mu = a) = P(T \notin [t_1, t_2]),$$

kjer sta

$$t_1 = \frac{x_1 - a}{s/\sqrt{n}} \text{ in } t_2 = \frac{x_2 - a}{s/\sqrt{n}}.$$

Če je vzorec velik (?), je slučajna spremenljivka $Z = \frac{\bar{X}-a}{s/\sqrt{n}}$ približno standardno normalno porazdeljena in je, na primer,

$$\alpha = P(\bar{X} \notin [x_1, x_2] \mid \mu = a) = P(Z \notin [z_1, z_2]),$$

kjer sta

$$z_1 = \frac{x_1 - a}{s/\sqrt{n}} \text{ in } z_2 = \frac{x_2 - a}{s/\sqrt{n}}.$$

Izračunajte α-napaki testiranja hipoteze H_0 na oba načina (normalna in t porazdelitev). Ali hipotezo H_0 zavrnemo?

j.) Izračunajte P-vrednost statističnega testa iz prejšnje točke (t-porazdelitev). Pomagate si lahko z ukazom

```
t.test(x, mu = 40, alternative = "two.sided"...).
```

k.) Določite interval zaupanja za povprečno vrednost x_p s stopnjo zaupanja $95\,\%$ (t-porazdelitev). Pomagate si lahko z ukazom

```
t.test(x, mu = 40, alternative = "two.sided", conf.level = 0.95, ...).
```

l.) Kaj lahko svetujemo tovarni?

R: a.) $n = 50$ b.) $xp = 40.61712$, $s = 2.173799$ c.) - d.) $H_0 : \mu = 40$, $H_1 : \mu \neq 40$ e.) $x_1 = 39.5$, $x_2 = 40.5$ f.) $\alpha = 0.07709987$, zavrnemo g.) $p = 0.02912175$ h.) $l = 40.063$, $u = 41.171$ i.) $\alpha = 0.1102729$ (t-porazdelitev), $\alpha = 0.103858$ (normalna porazdelitev), zavrnemo j.) $p = 0.0502397$ k.) $[39.999, 41.235]$ l.) -

3. Vsebino skriptne datoteke shranite pod imenom ‘vaja10.r’.