(B) Vzorčna disperzija

Imejmo normalno populacijo $N(\mu, \sigma)$.

Kako bi določili porazdelitev za vzorčno disperzijo $S_0^2 = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$

ali popravljeno vzorčno disperzijo $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$?

Raje izračunamo porazdelitev za statistiko

$$\chi^2 = \frac{nS_0^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

. . . Vzorčna disperzija

Preoblikujemo jo lahko takole:

$$\begin{aligned}\chi^2 &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu + \mu - \overline{X})^2 = \\ &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 + \frac{2}{\sigma^2}(\mu - \overline{X})\sum_{i=1}^{n}(X_i - \mu) + \frac{n}{\sigma^2}(\mu - \overline{X})^2\end{aligned}$$

in, ker je $\sum_{i=1}^{n}(X_i - \mu) = n(\overline{X} - \mu) = -n(\mu - \overline{X})$, dalje

$$\chi^2 = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 - \frac{1}{n}\left(\sum_{i=1}^{n}\frac{X_i - \mu}{\sigma}\right)^2 = \sum_{i=1}^{n}Y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n}Y_i\right)^2,$$

kjer so $Y_1, Y_2, \ldots, Y_n$ paroma neodvisne standardizirano normalno porazdeljene slučajne spremenljivke, $Y_i = \dfrac{X_i - \mu}{\sigma}$.

. . . Vzorčna disperzija

Porazdelitvena funkcija za χ^2 je

$$F_{\chi^2} = P(\chi^2 < z) = \int\!\!\int\cdots\int_{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}(\sum_{i=1}^{n} y_i)^2 < z} e^{-(y_1^2+y_2^2+\ldots+y_n^2)/2}dy_n \ldots dy_1$$

z ustrezno ortogonalno transformacijo v nove spremenljivke $z_1, z_2, \ldots, z_n$ dobimo po nekaj računanja

$$F_{\chi^2} = \frac{1}{(2\pi)^{(n-1)/2}}\int\!\!\int\cdots\int_{\sum_{i=1}^{n-1} z_i^2 < z} e^{-(z_1^2+z_2^2+\ldots+z_{n-1}^2)/2}dz_{n-1}\ldots dz_1$$

Pod integralom je gostota vektorja $(Z_1, Z_2, \ldots, Z_{n-1})$ z neodvisnimi standardizirano normalnimi členi. Integral sam pa ustreza porazdelitveni funkciji vsote kvadratov $Z_1^2 + Z_2^2 + \ldots + Z_{n-1}^2$.
Tako je porazdeljena tudi statistika χ^2.

. . . Vzorčna disperzija

Kakšna pa je ta porazdelitev? Ker so tudi kvadrati $Z_1^2, Z_2^2, \ldots, Z_{n-1}^2$ med seboj neodvisni in porazdeljeni po zakonu $\chi^2(1)$,
je njihova vsota porazdeljena po zakonu $\chi^2(n-1)$.
Tako je torej porazdeljena tudi statistika χ^2.

Ker vemo, da je $\mathsf{E}\chi^2(n) = n$ in $\mathsf{D}\chi^2(n) = 2n$, lahko takoj izračunamo

$$\mathsf{E}S_0^2 = \mathsf{E}\frac{\sigma^2\chi^2}{n} = \frac{(n-1)\sigma^2}{n} \qquad \mathsf{E}S^2 = \mathsf{E}\frac{\sigma^2\chi^2}{n-1} = \sigma^2$$

in

$$\mathsf{D}S_0^2 = \mathsf{D}\frac{\sigma^2\chi^2}{n} = \frac{2(n-1)\sigma^4}{n^2} \qquad \mathsf{D}S^2 = \mathsf{D}\frac{\sigma^2\chi^2}{n-1} = \frac{2\sigma^4}{n-1}$$

. . . Vzorčna disperzija

Če je n zelo velik, je po centralnem limitnem izreku statistika χ^2 porazdeljena približno normalno in sicer po zakonu

$$N\big(n-1, \sqrt{2(n-1)}\big),$$

vzorčna disperzija S_0^2 približno po

$$N\left(\frac{(n-1)\sigma^2}{n}, \frac{\sqrt{2(n-1)}\,\sigma^2}{n}\right)$$

in popravljena vzorčna disperzija S^2 približno po

$$N\left(\sigma^2, \sqrt{\frac{2}{n-1}}\,\sigma^2\right).$$

Studentova porazdelitev

Pri normalno porazdeljeni slučajni spremenljivki X

je tudi porazdelitev $\overline{X}$ normalna, in sicer $N(\mu, \dfrac{\sigma}{\sqrt{n}})$.

Statistika

$$Z = \frac{\overline{X} - \mu}{\sigma}\sqrt{n}$$

je potem porazdeljena standardizirano normalno.

Pri ocenjevanju parametra μ z vzorčnim povprečjem $\overline{X}$
to lahko uporabimo le, če poznamo σ;
sicer ne moremo oceniti standardne napake
– ne vemo, kako dobra je ocena za μ.

Kaj lahko naredimo, če σ ne poznamo?

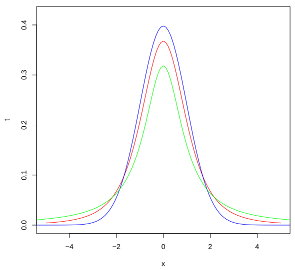
Parameter σ lahko ocenimo s S_0 ali S.
Toda S je slučajna spremenljivka in porazdelitev statistike $\dfrac{\overline{X}-\mu}{S}\sqrt{n}$ *ni več* normalna $N(0,1)$ (razen, če je n zelo velik in S skoraj enak σ).

Kakšna je porazdelitev nove vzorčne statistike

$$T = \frac{\overline{X}-\mu}{S}\sqrt{n}\ ?$$


'Student' in 1908

. . . Studentova porazdelitev



Leta 1908 je W.S. Gosset (1876-1937) pod psevdonimom 'Student' objavil članek, v katerem je pokazal, da ima statistika T porazdelitev $S(n-1)$ z gostoto

$$p(t) = \frac{\left(1+\dfrac{t^2}{n-1}\right)^{-\dfrac{n}{2}}}{\sqrt{n-1}\,B\left(\dfrac{n-1}{2},\dfrac{1}{2}\right)}$$

Tej porazdelitvi pravimo **Studentova porazdelitev** *z $n-1$ prostostnimi stopnjami.*

```
> plot(function(x) dt(x,df=3),-5,5,ylim=c(0,0.42),ylab="t",
  col="red")
> curve(dt(x,df=100),col="blue",add=T)
> curve(dt(x,df=1),col="green",add=T)
```


. . . Studentova porazdelitev

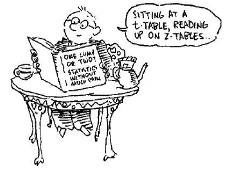
Za $S(1)$ dobimo Cauchyevo porazdelitev z gostoto

$$p(t) = \frac{1}{\pi(1+t^2)}$$

Za $n\to\infty$ pa gre $\frac{1}{\sqrt{n-1}\,B(\frac{n-1}{2},\frac{1}{2})} \to \sqrt{2\pi}$ in $(1+\frac{t^2}{n-1})^{-\frac{n}{2}} \to e^{-\frac{t^2}{2}}$.

Torej ima limitna porazdelitev gostoto

$$p(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$$



standardizirane normalne porazdelitve.

Če zadnji sliki dodamo

```
> curve(dnorm(x),col="magenta",add=T)
```

ta pokrije modro krivuljo.

Fisherjeva ali Snedecorjeva porazdelitev

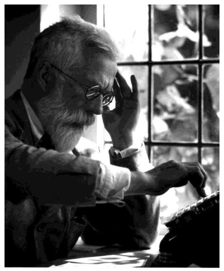
Poskusimo najti še porazdelitev kvocienta $Z=\dfrac{U}{V}$,

kjer sta $U:\chi^2(m)$ in $V:\chi^2(n)$ ter sta U in V neodvisni.

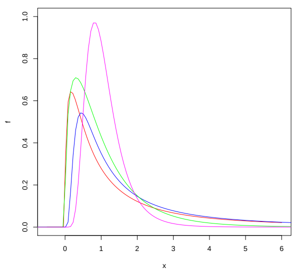
Z nekaj računanja (glej Hladnik) je mogoče pokazati, da je za $x>0$ gostota ustrezne porazdelitve $F(m,n)$ enaka

$$p(x) = \frac{m^{\frac{m}{2}}n^{\frac{n}{2}}}{B(\frac{m}{2},\frac{n}{2})}\,\frac{x^{\frac{m}{2}-1}}{(n+mx)^{\frac{m+n}{2}}}$$

in je enaka 0 drugje.



. . . Fisherjeva porazdelitev



Porazdelitvi $F(m,n)$ pravimo **Fisherjeva** ali tudi **Snedecorjeva porazdelitev** ***F*** *z (m,n) prostostnimi stopnjami.*

```
> plot(function(x) df(x,df1=3,df2=2),-0.5,6,ylim=c(0,1),ylab="f",
  col="red")
> curve(df(x,df1=20,df2=2),col="blue",add=T)
> curve(df(x,df1=3,df2=20),col="green",add=T)
> curve(df(x,df1=20,df2=20),col="magenta",add=T)
```


. . . Fisherjeva porazdelitev

Po zakonu $F(m-1,n-1)$ je na primer porazdeljena statistika

$$F = \frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2}$$

saj vemo, da sta spremenljivki

$$U=(m-1)S_X^2/\sigma_X^2 \quad \text{in} \quad V=(n-1)S_Y^2/\sigma_Y^2$$

porazdeljeni po χ^2 z $m-1$ oziroma $n-1$ prostostnimi stopnjami in sta neodvisni.

Velja še:

če je $U:F(m,n)$, je $1/U:F(n,m)$,

če je $U:S(n)$, je $U^2:F(1,n)$.

Intervalno ocenjevanje in cenilke

Merjenje slučajne spremenljivke

Vzorec: $(X_1, X_2, \ldots, X_n)$, kjer je n velikost, je slučajni vektor.

Če imajo komponente X_i enako porazdelitev in so neodvisne, rečemo, da gre za **enostavni slučajni vzorec**.

Lastnosti vzorčnega povprečja

1. Če je na preučevani populaciji slučajna spremenljivka porazdeljena normalno, je tako porazdeljeno tudi vzorčno povprečje.

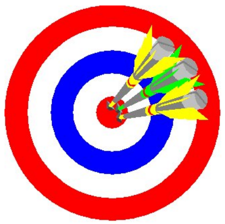
2. Vzorčno povprečje se pri velikih vzorčih porazdeljuje **približno normalno** tudi, če je spremenljivka na osnovni populaciji porazdeljena kako drugače.

3. Matematični upanji spremenljivk X in $\overline{X}$ sta enaki.

Točkovne cenilke

Točkovna cenilka je pravilo ali formula, ki nam pove, kako izračunati numerično cenilko na osnovi merjenj vzorca.

Število, ki je rezultat izračuna, se imenuje **točkovna cenilka**.



Cenilke

Cenilka parametra ζ je vzorčna statistika $C = C(X_1, X_2, X_3, \ldots, X_n)$, katere porazdelitveni zakon je odvisen le od parametra ζ, njene vrednosti pa ležijo v prostoru parametrov.

Od cenilke običajno pričakujemo, da je simetrična – njena vrednost je enaka za vse permutacije argumentov. Seveda je odvisna tudi od velikosti vzorca n .

Primeri: vzorčna mediana $\tilde{X}$ in vzorčno povprečje $\overline{X}$ sta cenilki za populacijsko povprečje μ;

popravljena vzorčna disperzija S^2 pa je cenilka za populacijsko disperzijo σ^2 .

Doslednost

Cenilka C parametra ζ je **dosledna**, če z rastočim n zaporedje C_n verjetnostno konvergira k ζ, to je, za vsak $\varepsilon > 0$ velja

$$\lim_{n\to\infty} P(|C_n - \zeta| < \varepsilon) = 1$$

Primeri: vzorčno povprečje $\overline{X}$ je dosledna cenilka za populacijsko povprečje μ. Tudi vsi **vzorčni začetni momenti**

$$Z_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

so dosledne cenilke ustreznih začetnih populacijskih momentov $z_k = \mathsf{E}X^k$, če le-ti obstajajo.

Vzorčna mediana $\tilde{X}$ je dosledna cenilka za populacijsko mediano.

. . . Doslednost

Če pri pogoju $n \to \infty$ velja $\mathsf{E}C_n \to \zeta$ in $\mathsf{D}C_n \to 0$, je C_n dosledna cenilka parametra ζ.

To sprevidimo takole:

$$1 - P(|C_n - \zeta| < \varepsilon) = P(|C_n - \zeta| \geq \varepsilon) \leq P(|C_n - \mathsf{E}C_n| + |\mathsf{E}C_n - \zeta| \geq \varepsilon)$$

upoštevajmo še, da za dovolj velike n velja $|\mathsf{E}C_n - \zeta| < \varepsilon/2$, in uporabimo neenakost Čebiševa

$$P(|C_n - \mathsf{E}C_n| \geq \epsilon/2) \leq \frac{4\mathsf{D}C_n}{\epsilon^2} \to 0$$

Primeri: Naj bo $X : N(\mu, \sigma)$. Ker za $n \to \infty$ velja $\mathsf{E}S_0^2 = \frac{(n-1)\sigma^2}{n} \to \sigma^2$ in $\mathsf{D}S_0^2 = \frac{2(n-1)\sigma^4}{n^2} \to 0$, je vzorčna disperzija S_0^2 dosledna cenilka za σ^2.

Nepristrana cenilka z najmanjšo varianco

Cenilka C_n parametra ζ je **nepristranska**, če je $\mathsf{E}C_n = \zeta$ (za vsak n); in je **asimptotično nepristranska**, če je $\lim_{n\to\infty} \mathsf{E}C_n = \zeta$.

Količino $B(C_n) = \mathsf{E}C_n - \zeta$ imenujemo **pristranost** (angl. *bias*) cenilke C_n.

Primeri: vzorčno povprečje $\overline{X}$ je nepristranska cenilka za populacijsko povprečje μ ;

vzorčna disperzija S_0^2 je samo asimptotično nepristranska cenilka za σ^2, popravljena vzorčna disperzija S^2 pa je nepristranska cenilka za σ^2.

Disperzija nepristranskih cenilk

Izmed nepristranskih cenilk istega parametra ζ je boljša tista, ki ima manjšo disperzijo – v povprečju daje bolj točne ocene.

Če je razred cenilk parametra ζ *konveksen* (vsebuje tudi njihove konveksne kombinacije), obstaja v bistvu ena sama cenilka z najmanjšo disperzijo:

Naj bo razred nepristranskih cenilk parametra ζ konveksen. Če sta C in C' nepristranski cenilki, obe z najmanjšo disperzijo σ^2, je $C = C'$ z verjetnostjo 1.

Za to poglejmo

$$\mathsf{D}(\frac{1}{2}(C+C')) = \frac{1}{4}(\mathsf{D}C + \mathsf{D}C' + 2\mathsf{Cov}(C,C')) \leq (\frac{1}{2}(\sqrt{\mathsf{D}C} + \sqrt{\mathsf{D}C'}))^2 = \sigma^2$$

Ker sta cenilki minimalni, mora biti tudi $\mathsf{D}(\frac{1}{2}(C + C')) = \sigma^2$ in dalje $\mathsf{Cov}(C, C') = \sigma^2$ oziroma $r(C, C') = 1$. Torej je $C' = aC + b$, $a > 0$ z verjetnostjo 1. Iz $\mathsf{D}C = \mathsf{D}C'$ izhaja $a = 1$, iz $\mathsf{E}C = \mathsf{E}C'$ pa še $b = 0$.

Srednja kvadratična napaka

Včasih je celo bolje vzeti pristransko cenilko z manjšo disperzijo, kot jo ima druga, sicer nepristranska, cenilka z veliko disperzijo.

Mera *učinkovitosti* cenilk parametra ζ je *srednja kvadratična napaka*

$$q(C) = \mathsf{E}(C - \zeta)^2$$

Ker velja

$$q(C) = \mathsf{E}(C - \mathsf{E}C + \mathsf{E}C - \zeta)^2 = \mathsf{E}(C - \mathsf{E}C)^2 + (\mathsf{E}C - \zeta)^2$$

jo lahko zapišemo tudi v obliki

$$q(C) = \mathsf{D}C + B(C)^2$$

Za nepristranske cenilke je $B(C) = 0$ in zato $q(C) = \mathsf{D}C$.

Če pa je disperzija cenilke skoraj 0, je $q(C) \approx B(C)^2$.

Rao-Cramérjeva ocena

Naj bo f gostotna ali verjetnostna funkcija slučajne spremenljivke X in naj bo odvisna še od parametra ζ, tako da je $f(x;\zeta)$ njena vrednost v točki x. Združeno gostotno ali verjetnostno funkcijo slučajnega vzorca $(X_1, X_2, X_3, \ldots, X_n)$ označimo z L in ji pravimo *funkcija verjetja* (tudi *zanesljivosti*, angl. *likelihood*)

$$L(x_1, x_2, x_3, \ldots, x_n; \zeta) = f(x_1;\zeta) f(x_2;\zeta) f(x_3;\zeta) \cdots f(x_n;\zeta)$$

Velja (*): $\int\int \ldots \int L(x_1, x_2, \ldots, x_n; \zeta) dx_1 dx_2 \ldots dx_n = 1$.

$L(X_1, X_2, X_3, \ldots, X_n)$ je funkcija vzorca – torej slučajna spremenljivka.

Privzemimo, da je funkcija L vsaj dvakrat zvezno odvedljiva po ζ na nekem intervalu I in naj na tem intervalu tudi integral odvoda L po ζ enakomerno konvergira.

. . . Rao-Cramérjeva ocena

Odvajajmo enakost (*) po ζ in upoštevajmo $\frac{\partial \ln L}{\partial \zeta} = \frac{1}{L}\frac{\partial L}{\partial \zeta}$ pa dobimo

$$\int\int \ldots \int \frac{\partial \ln L}{\partial \zeta} L dx_1 dx_2 \ldots dx_n = 0$$

kar lahko tolmačimo kot $\mathsf{E}\frac{\partial \ln L}{\partial \zeta} = 0$.

Naj bo sedaj C nepristranska cenilka parametra ζ, torej $\mathsf{E}C = \zeta$, oziroma zapisano z integrali $\int\int \ldots \int C L dx_1 dx_2 \ldots dx_n = \zeta$.

Ker C ni odvisna od ζ, dobimo z odvajanjem po ζ:

$$\int\int \ldots \int C \frac{\partial \ln L}{\partial \zeta} L dx_1 dx_2 \ldots dx_n = 1$$

kar pomeni $\mathsf{E}(C\frac{\partial \ln L}{\partial \zeta}) = 1$.

. . . Rao-Cramérjeva ocena

Če to enakost združimo s prejšnjo (pomnoženo s ζ), dobimo:

$$\mathsf{E}\left((C-\zeta)\,\frac{\partial \ln L}{\partial \zeta}\right) = 1$$

Od tu po $(\mathsf{E}XY)^2 \le \mathsf{E}X^2\mathsf{E}Y^2$ izhaja naprej

$$1 = \left(\mathsf{E}\left((C-\zeta)\,\frac{\partial \ln L}{\partial \zeta}\right)\right)^2 \le \mathsf{E}(C-\zeta)^2\,\mathsf{E}\left(\frac{\partial \ln L}{\partial \zeta}\right)^2 = \mathsf{D}C\;\mathsf{E}\left(\frac{\partial \ln L}{\partial \zeta}\right)^2$$

kar da *Rao-Cramérjevo oceno*

$$\mathsf{D}C \ge \left(\mathsf{E}\left(\frac{\partial \ln L}{\partial \zeta}\right)^2\right)^{-1} = \left(-\mathsf{E}\frac{\partial^2 \ln L}{\partial \zeta^2}\right)^{-1} = \left(n\mathsf{E}\left(\frac{\partial \ln f}{\partial \zeta}\right)^2\right)^{-1}$$

Učinkovitost cenilk

Rao-Cramérjeva ocena da absolutno spodnjo mejo disperzije za vse nepristranske cenilke parametra ζ (v dovolj gladkih porazdelitvah).
Ta meja ni nujno dosežena. Cenilka, ki jo doseže, se imenuje *najučinkivitejša cenilka* parametra ζ in je ena sama (z verjetnostjo 1).

Kdaj pa je ta spodnja meja dosežena?

V neenakosti $(\mathsf{E}XY)^2 \le \mathsf{E}X^2\mathsf{E}Y^2$, ki je uporabljena v izpeljavi Rao-Cramérjeve ocene, velja enakost natanko takrat, ko je $Y = cX$ z verjetnostjo 1.

. . . Učinkovitost cenilk

Torej velja v Rao-Cramérjevi oceni enakost natanko takrat, ko je

$$\frac{\partial \ln L}{\partial \zeta} = A(\zeta)(C-\zeta)$$

kjer je $A(\zeta)$ konstanta, odvisna od ζ in neodvisna od vzorca.

Zato je tudi

$$(\mathsf{D}C)^{-1} = \mathsf{E}\left(\frac{\partial \ln L}{\partial \zeta}\right)^2 = A(\zeta)^2\mathsf{E}(C-\zeta)^2 = A(\zeta)^2\mathsf{D}C$$

oziroma končno

$$\mathsf{D}C = |A(\zeta)|^{-1}$$

Najučinkovitejše cenilke za parametre normalne porazdelitve

Naj bo $X : N(\mu,\sigma)$. Tedaj je

$$L = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-((\frac{X_1-\mu}{\sigma})^2+\cdots+(\frac{X_n-\mu}{\sigma})^2)/2}$$

in

$$\ln L = \ln \frac{1}{(2\pi)^{n/2}\sigma^n} - ((\frac{X_1-\mu}{\sigma})^2+\cdots+(\frac{X_n-\mu}{\sigma})^2)/2$$

ter dalje

$$\frac{\partial \ln L}{\partial \mu} = \frac{X_1-\mu}{\sigma^2}+\cdots+\frac{X_n-\mu}{\sigma^2} = \frac{n}{\sigma^2}(\overline{X}-\mu)$$

Torej je vzorčno povprečje $\overline{X}$ najučinkovitejša cenilka za μ z disperzijo $\mathsf{D}\overline{X} = \frac{\sigma^2}{n}$.

. . . normalna porazdelitev

Prvi člen v izrazu za $\ln L$ lahko zapišemo tudi $-\frac{n}{2}(\ln 2\pi + \ln \sigma^2)$. Tedaj je, če privzamemo, da je μ znano število

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}((X_1-\mu)^2+\cdots+(X_n-\mu)^2) = \frac{n}{2\sigma^4}(S_\mu^2-\sigma^2)$$

To pomeni, da je $S_\mu^2 = \frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2$ najučinkovitejša cenilka za parameter σ^2 z disperzijo $\mathsf{D}S_\mu^2 = \frac{2\sigma^4}{n}$.

Poissonova porazdelitev

Za Poissonovo porazdelitev $P(\lambda)$ s parametrom λ, $p_k = \lambda^k \frac{e^{-\lambda}}{k!}$ je

$$L = e^{-n\lambda}\frac{\lambda^{x_1+\cdots+x_n}}{x_1!\cdots x_n!}$$

in dalje

$$\ln L = -n\lambda + (x_1+\cdots+x_n)\ln\lambda - \ln(x_1!\cdots x_n!)$$

ter končno

$$\frac{\partial \ln L}{\partial \lambda} = -n + \frac{x_1+\cdots+x_n}{\lambda} = \frac{n}{\lambda}(\overline{X}-\lambda)$$

Najučinkovitejša cenilka za parameter λ je $\overline{X}$ z disperzijo $\mathsf{D}\overline{X} = \frac{\lambda}{n}$.

Učinkovitost cenilke

Naj bo C_0 najučinkovitejša cenilka parametra ζ in C kaka druga nepristranska cenilka. Tedaj je *učinkovitost* cenilke C določena s predpisom

$$e(C) = \frac{\mathsf{D}C_0}{\mathsf{D}C}$$

Učinkovitost najučinkovitejše cenilke je $e(C_0) = 1$.

Če najučinkovitejša cenilka ne obstaja, vzamemo za vrednost $\mathsf{D}C_0$ desno stran v Rao-Cramérjevi oceni.

Primer: Naj bo $X : N(\mu, \sigma)$. Pri velikih n-jih je vzorčna mediana $\tilde{X}$ – ocena za μ, porazdeljena približno po $N(\mu, \sigma\sqrt{\frac{\pi}{2n}})$. Torej je

$$e(\tilde{X}) = \frac{\mathsf{D}\overline{X}}{\mathsf{D}\tilde{X}} = \frac{\frac{\sigma^2}{n}}{\frac{\pi\sigma^2}{2n}} = \frac{2}{\pi} \approx 0.64$$

...Učinkovitost cenilke

Primer: Naj bo $X : N(\mu, \sigma)$. Če poznamo μ, je najučinkovitejša cenilka za σ^2 statistika S^2_μ z disperzijo $\mathsf{D}S^2_\mu = \frac{2\sigma^4}{n}$. Popravljena vzorčna disperzija S^2 pa je nepristranska cenilka istega parametra z disperzijo $\mathsf{D}S^2 = \frac{2\sigma^4}{n-1}$. Torej je učinkovitost S^2

$$e(S^2) = \frac{\mathsf{D}S^2_\mu}{\mathsf{D}S^2} = \frac{\frac{2\sigma^4}{n}}{\frac{2\sigma^4}{n-1}} = \frac{n-1}{n}$$

Iz tega vidimo, da $e(S^2) \to 1$, ko $n \to \infty$. Pravimo, da je cenilka S^2 *asimptotično najučinkovitejša cenilka* za σ^2.

Metoda momentov

Recimo, da je za zvezno slučajno spremenljivko X njena gostota f odvisna od m parametrov $f(x; \zeta_1, \zeta_2, \zeta_3, \dots, \zeta_m)$ in naj obstajajo momenti

$$z_k = z_k(\zeta_1, \zeta_2, \zeta_3, \dots, \zeta_m) = \int_{-\infty}^{\infty} x^k f(x; \zeta_1, \zeta_2, \zeta_3, \dots, \zeta_m)dx$$

za $k = 1, 2, 3, \dots, m$. Če se dajo iz teh enačb enolično izračunati parametri $\zeta_1, \zeta_2, \zeta_3, \dots, \zeta_m$ kot funkcije momentov $z_1, z_2, z_3, \dots, z_m$

$$\zeta_k = \varphi_k(z_1, z_2, z_3, \dots, z_m)$$

potem so

$$C_k = \varphi_k(Z_1, Z_2, Z_3, \dots, Z_m)$$

cenilke parametrov ζ_k po *metodi momentov*. k-ti vzorčni začetni moment $Z_k = \frac{1}{n}\sum_{i=1}^n X_i^k$ je cenilka za ustrezni populacijski moment z_k.

Cenilke, ki jih dobimo po metodi momentov so dosledne.

...Metoda momentov

Naj bo $X : N(\mu, \sigma)$. Tedaj je $z_1 = \mu$ in $z_2 = \sigma^2 + \mu^2$.

Od tu dobimo $\mu = z_1$ in $\sigma^2 = z_2 - z_1^2$.

Ustrezni cenilki sta $Z_1 = \overline{X}$ za μ in

$$Z_2 - Z_1^2 = \overline{X^2} - \overline{X}^2 = S_0^2$$

za σ^2 – torej vzorčno povprečje in disperzija.

Metoda največjega verjetja

Funkcija verjetja

$$L(x_1, x_2, x_3, \dots, x_n; \zeta) = f(x_1; \zeta)f(x_2; \zeta)f(x_3; \zeta)\cdots f(x_n; \zeta)$$

je pri danih $x_1, x_2, x_3, \dots, x_n$ odvisna še od parametra ζ. Izberemo tak ζ, da bo funkcija L dosegla največjo vrednost. Če je L vsaj dvakrat zvezno odvedljiva, mora veljati $\frac{\partial L}{\partial \zeta} = 0$ in $\frac{\partial^2 L}{\partial \zeta^2} < 0$. Največja vrednost parametra je še odvisna od $x_1, x_2, x_3, \dots, x_n$:
$\zeta_{max} = \varphi(x_1, x_2, x_3, \dots, x_n)$. Tedaj je cenilka za parameter ζ enaka

$$C = \varphi(X_1, X_2, X_3, \dots, X_n)$$

Metodo lahko posplošimo na večje število parametrov.

Pogosto raje iščemo maksimum funkcije $\ln L$.

Če najučinkovitejša cenilka obstaja, jo dobimo s to metodo.

...Metoda največjega verjetja - binomska

Naj bo $X : B(1, p)$. tedaj je $f(x; p) = p^x(1-p)^{1-x}$, kjer je $x = 0$ ali $x = 1$. Ocenjujemo parameter p. Funkcija verjetja ima obliko $L = p^x(1-p)^{n-x}$, kjer je sedaj $x \in \{0, 1, 2, \dots, n\}$. Ker je $\ln L = x\ln p + (n-x)\ln(1-p)$, dobimo

$$\frac{\partial \ln L}{\partial p} = \frac{x}{p} - \frac{n-x}{1-p},$$

ki je enak 0 pri $p = \frac{x}{n}$. Ker je v tem primeru $\frac{\partial^2 \ln L}{\partial p^2} = -\frac{x}{p^2} - \frac{n-x}{(1-p)^2} < 0$, je v tej točki maksimum. Cenilka po metodi največjega verjetja je torej $P = \frac{X}{n}$, kjer je X binomsko porazdeljena spremenljivka – frekvenca v n ponovitvah. Cenilka P je nepristranska, saj je $\mathsf{E}P = \frac{\mathsf{E}X}{n} = p$. Ker za $n \to \infty$ gre $\mathsf{D}P = \frac{\mathsf{D}X}{n^2} = \frac{p(1-p)}{n} \to 0$, je P dosledna cenilka. P je tudi najučinkovitejša $\frac{\partial \ln L}{\partial p} = \frac{X}{p} - \frac{n-X}{1-p} = \frac{n}{p(1-p)}(\frac{X}{n} - p) = \frac{n}{p(1-p)}(P - p)$.

. . . Metoda največjega verjetja - Poissonova

Za Poissonovo porazdelitev $P(\lambda)$ s parametrom λ, $p_x = \lambda^x \frac{e^{-\lambda}}{x!}$ je

$$L = e^{-n\lambda} \frac{\lambda^{x_1+\cdots+x_n}}{x_1! \cdots x_n!}$$

in dalje

$$\ln L = -n\lambda + (x_1 + \cdots + x_n) \ln \lambda - \ln(x_1! \cdots x_n!)$$

ter končno

$$\frac{\partial \ln L}{\partial \lambda} = -n + \frac{x_1 + \cdots + x_n}{\lambda} = \frac{n}{\lambda}(\overline{X} - \lambda)$$

Odvod je enak 0 za $\lambda = \overline{X}$. Drugi odvod v tej točki je $\frac{\partial^2 \ln L}{\partial \lambda^2} = -\frac{x_1+\cdots+x_n}{\lambda^2} < 0$. V točki je maksimum.

Cenilka za λ po metodi največjega verjetja je vzorčno povprečje $\overline{X}$.

Je tudi najučinkovitejša cenilka za λ z disperzijo $\mathsf{D}\overline{X} = \frac{\lambda}{n}$.

Porazdelitev vzorčnih aritmetičnih sredin

Primer: Denimo, da se spremenljivka inteligenčni kvocient na populaciji porazdeljuje normalno z aritmetično sredino $\mu = 100$ in standardnim odklonom $\sigma = 15$.

$$X : N(100, 15)$$

Denimo, da imamo vzorec velikosti $n = 225$. Tedaj se vzorčne aritmetične sredine porazdeljujejo normalno

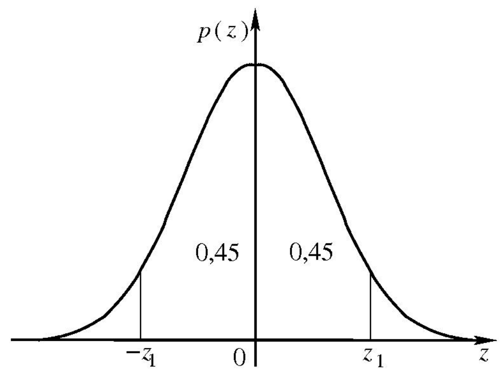
$$\overline{X} : N(100, \frac{15}{\sqrt{225}}) = N(100, 1)$$

Izračunajmo, kolikšne vzorčne aritmetične sredine ima 90% vzorcev (simetrično na povprečje). 90% vzorčnih aritmetičnih sredin se nahaja na intervalu:

$$P(\overline{X}_1 < \overline{X} < \overline{X}_2) = 0,90$$

$$P(-z_1 < z < z_1) = 0,90 \quad \Longrightarrow \quad 2\Phi(z_1) = 0,90$$

$$\Phi(z_1) = 0,45 \quad \Longrightarrow \quad z_1 = 1,65$$



Potem se vzorčne aritmetične sredine nahajajo v intervalu

$$P\left(\mu - z_1 \frac{\sigma}{\sqrt{n}} < \overline{X} < \mu + z_1 \frac{\sigma}{\sqrt{n}}\right) = 0,90$$

oziroma konkretno

$$P\left(100 - 1,65 < \overline{X} < 100 + 1,65\right) = 0,90$$

90% vseh slučajnih vzorcev velikosti 225 enot bo imelo povprečja za inteligenčni kvocient na intervalu

$$(98,35\,;\,101,65).$$

Lahko preverimo, da bi bil ta interval v primeru večjega vzorca ožji. Npr. v primeru vzorcev velikosti $n = 2500$ je ta interval

$$P\left(100 - 1,65 \frac{15}{\sqrt{2500}} < \overline{X} < 100 + 1,65 \frac{15}{\sqrt{2500}}\right) = 0,90$$

oziroma

$$(99,5\,;\,100,5).$$

Porazdelitev vzorčnih deležev

Denimo, da želimo na populaciji oceniti delež enot π z določeno lastnostjo.

Zato na vsakem vzorcu poiščemo vzorčni delež p.

Pokazati se da, da se za dovolj velike slučajne vzorce s ponavljanjem (za deleže okoli 0,5 je dovolj 20 enot ali več) vzorčni deleži porazdeljujejo približno normalno z

- aritmetično sredino vzorčnih deležev, ki je enaka deležu na populaciji

$$\mathsf{E}p = \pi,$$

- standardnim odklonom vzorčnih deležev

$$\mathsf{SE}(p) = \sqrt{\frac{\pi(1-\pi)}{n}}.$$

Za manjše vzorce se vzorčni deleži porazdeljujejo binomsko.

Cenilka populacijskega deleža je nepristranska cenilka, ker velja

$$\mathsf{E}p = \pi.$$

Primer: V izbrani populaciji prebivalcev je polovica žensk $\pi = 0,5$. Če tvorimo vzorce po $n = 25$ enot, nas zanima, kolikšna je verjetnost, da je v vzorcu več kot 55 % žensk? To pomeni, da iščemo verjetnost $P(p > 0,55)$.

Vzorčni deleži p se porazdeljujejo približno normalno

$$p \ : \ N\left(0,5\ ,\ \sqrt{\frac{\pi(1-\pi)}{n}}\right) = N\left(0,5\ ,\ \sqrt{\frac{0,5\cdot 0,5}{25}}\right) = N(0,5\ ,\ 0,1).$$

$$P(p > 0,55) = P\left(Z > \frac{0,55-0,5}{0,1}\right) = P(Z > 0,5) =$$

$$= 0,5 - \Phi(0,5) = 0,5 - 0,1915 = 0,3085.$$

Rezultat pomeni, da lahko pričakujemo, da bo pri približno 31% vzorcev delež žensk večji od 0,55.

Poglejmo, kolikšna je ta verjetnost, če bi tvorili vzorce velikosti $n = 2500$ enot:

$$P(p > 0,55) = P\left(Z > \frac{0,55-0,5}{\sqrt{\frac{0,5(1-0,5)}{2500}}}\right)$$

$$= P(Z > 5) = 0,5 - \Phi(5) = 0,5 - 0,5 = 0.$$

V 10-krat večjih vzorcih kot prej ne moremo pričakovati več kot 55% žensk.

Porazdelitev razlik vzorčnih aritmetičnih sredin

Denimo, da imamo dve populaciji velikosti N_1 in N_2 in se spremenljivka X na prvi populaciji porazdeljuje normalno $N(\mu_1, \sigma)$, na drugi populaciji pa $N(\mu_2, \sigma)$ (standardna odklona sta na obeh populacijah enaka!).

V vsaki od obeh populacij tvorimo neodvisno slučajne vzorce velikosti n_1 in n_2. Na vsakem vzorcu (s ponavljanjem) prve populacije izračunamo vzorčno aritmetično sredino $\overline{X}_1$ in podobno na vsakem vzorcu druge populacije $\overline{X}_2$.

Dokazati se da, da je porazdelitev razlik vzorčnih aritmetičnih sredin normalna, kjer je

- matematično upanje razlik vzorčnih aritmetičnih sredin enako

$$\mathsf{E}(\overline{X}_1 - \overline{X}_2) = \mathsf{E}\overline{X}_1 - \mathsf{E}\overline{X}_2 = \mu_1 - \mu_2,$$

- disperzija razlik vzorčnih aritmetičnih sredin enaka

$$\mathsf{D}(\overline{X}_1 - \overline{X}_2) = \mathsf{D}\overline{X}_1 + \mathsf{D}\overline{X}_2 = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} = \sigma^2 \cdot \frac{n_1+n_2}{n_1 n_2}.$$

Primer: Dvema populacijama študentov na neki univerzi (tehnikom in družboslovcem) so izmerili neko sposobnost s povprečjema $\mu_t = 70$ in $\mu_d = 80$ točk in standardnim odklonom, ki je na obeh populacijah enak, $\sigma = 7$ točk.

Kolikšna je verjetnost, da je aritmetična sredina slučajnega vzorca družboslovcev ($n_d = 36$) večja za več kot 12 točk od aritmetične sredine vzorca tehnikov ($n_t = 64$)? Zanima nas torej verjetnost:

$$P(\overline{X}_d - \overline{X}_t > 12) = P\left(Z > \frac{12-10}{7\sqrt{\frac{36+64}{36\cdot 64}}}\right)$$

$$= P(Z > 1,37) = 0,5 - \Phi(1,37) =$$

$$= 0,5 - 0,4147 = 0,0853.$$

Torej, približno 8,5% parov vzorcev je takih, da je povprečje družboslovcev večje od povprečja tehnikov za 12 točk.

Porazdelitev razlik vzorčnih deležev

Podobno kot pri porazdelitvi razlik vzorčnih aritmetičnih sredin naj bosta dani dve populaciji velikosti N_1 in N_2 z deležema enot z neko lastnostjo π_1 in π_2. Iz prve populacije tvorimo slučajne vzorce velikosti n_1 in na vsakem izračunamo delež enot s to lastnostjo p_1.

Podobno naredimo tudi na drugi populaciji: tvorimo slučajne vzorce velikosti n_2 in na njih dolocimo deleže p_2. Pokazati se da, da se za dovolj velike vzorce razlike vzorčnih deležev porazdeljujejo približno normalno z

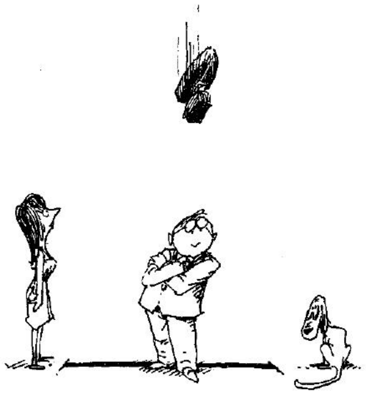
- matematičnim upanjem razlik vzorčnih deležev

$$\mathsf{E}(p_1 - p_2) = \mathsf{E}p_1 - \mathsf{E}p_2 = \pi_1 - \pi_2,$$

- disperzijo razlik vzorčnih deležev

$$\mathsf{D}(p_1 - p_2) = \mathsf{D}p_1 + \mathsf{D}p_2 = \frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}.$$

II.4. Intervali zaupanja



Denimo, da s slučajnim vzorcem ocenjujemo parameter γ.
Poskušamo najti statistiko g, ki je nepristranska, tj. $\mathsf{E}g = \gamma$ in se na vseh možnih vzorcih vsaj približno normalno porazdeljuje s standardno napako $\mathsf{SE}(g)$. Nato poskušamo najti interval, v katerem se bo z dano gotovostjo $(1-\alpha)$ nahajal ocenjevani parameter:

$$P(a < \gamma < b) = 1 - \alpha$$

a je spodnja meja zaupanja, b je zgornja meja zaupanja,
α verjetnost tveganja oziroma $1-\alpha$ verjetnost gotovosti.

Ta interval imenujemo **interval zaupanja** in ga interpretiramo takole:
z verjetnostjo tveganja α se parameter γ nahaja v tem intervalu.

Konstruirajmo interval zaupanja.

Na osnovi omenjenih predpostavk o porazdelitvi statistike g lahko zapišemo, da se statistika

$$Z = \frac{g - \mathsf{E}g}{\mathsf{SE}(g)} = \frac{g-\gamma}{\mathsf{SE}(g)}$$

porazdeljuje standardizirano normalno $N(0,1)$.

Če tveganje α porazdelimo polovico na levo in polovico na desno na konce normalne porazdelitve, lahko zapišemo

$$P\left(-z_{\alpha/2} < \frac{g-\gamma}{\mathsf{SE}(g)} < z_{\alpha/2}\right) = 1 - \alpha.$$

Po ustrezni preureditvi lahko izpeljemo naslednji interval zaupanja za parameter γ

$$P\Big(g - z_{\alpha/2}\,\mathsf{SE}(g) < \ \gamma \ < g + z_{\alpha/2}\,\mathsf{SE}(g)\Big) = 1 - \alpha$$

$z_{\alpha/2}$ je določen le s stopnjo tveganja α.

Vrednosti $z_{\alpha/2}$ lahko razberemo iz tabele za verjetnosti za standardizirano normalno porazdelitev, ker velja

$$\Phi(z_{\alpha/2}) = 0,5 - \frac{\alpha}{2}$$

Podajmo vrednost $z_{\alpha/2}$ za nekaj najbolj standardnih tveganj:

- $\alpha = 0,10, \quad z_{\alpha/2} = 1,65$
- $\alpha = 0,05, \quad z_{\alpha/2} = 1,96$
- $\alpha = 0,01, \quad z_{\alpha/2} = 2,58$

Pomen stopnje tveganja pri intervalih zaupanja

Za vsak slučajni vzorec lahko ob omenjenih predpostavkah izračunamo ob izbrani stopnji tveganja α interval zaupanja za parameter γ.

Ker se podatki vzorcev razlikujejo, se razlikujejo vzorčne ocene parametrov in zato tudi izračunani intervali zaupanja za parameter γ.
To pomeni, da se intervali zaupanja od vzorca do vzorca razlikujejo.
Meji intervala sta slučajni spremenljivki.

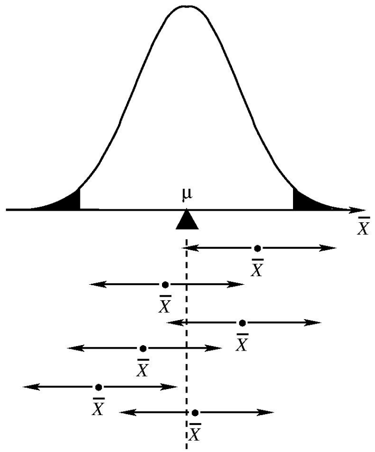
Vzemimo stopnjo tveganja $\alpha = 0,05$. Denimo, da smo izbrali 100 slučajnih vzorcev in za vsakega izračunali interval zaupanja za parameter γ.

Tedaj lahko pričakujemo, da 5 intervalov zaupanja od 100 ne bo pokrilo iskanega parametra γ.

Povedano je lepo grafično predstavljeno na naslednji strani.
V tem primeru ocenjujemo parameter aritmetično sredino inteligenčnega kvocienta. Kot vemo, se vzorčne aritmetične sredine $\overline{X}$ za dovolj velike vzorce porazdeljujejo nomalno.

Denimo, da v tem primeru poznamo vrednost parametra ($\mu = 100$).
Za več slučajnih vzorcev smo izračunali in prikazali interval zaupanja za μ ob stopnji tveganja $\alpha = 0,05$.

Predstavitev več intervalov zaupanja za aritmetično sredino μ pri 5% stopnji tveganja: približno 95% intervalov pokrije parameter μ.



Intervalsko ocenjevanje parametrov

Naj bo X slučajna spremenljivka na populaciji G z gostoto verjetnosti odvisno od parametra ζ.

Slučajna množica $M \subset \mathbb{R}$, ki je odvisna le od slučajnega vzorca, ne pa od parametra ζ, se imenuje *množica zaupanja* za parameter ζ, če obstaja tako število α, $0 < \alpha < 1$, da velja $P(\zeta \in M) = 1 - \alpha$. Število $1-\alpha$ imenujemo tedaj *stopnja zaupanja*; število α pa *stopnja tveganja*.

Stopnja zaupanja je običajno 95% ali 99% – $\alpha = 0,05$ ali $\alpha = 0,01$.
Pove nam, kakšna je verjetnost, da M vsebuje vrednost parametra ζ ne glede na to, kakšna je njegova dejanska vrednost.

Če je množica M interval $M = [A, B]$, ji rečemo *interval zaupanja* (za parameter ζ).
Njegovi krajišči sta funkciji slučajnega vzorca – torej statistiki.

...Intervalsko ocenjevanje parametrov

Naj bo $X : N(\mu, \sigma)$ in recimo, da poznamo parameter σ in ocenjujemo parameter μ. Izberimo konstanti a in b, $b > a$, tako da bo $P(a \le Z \le b) = 1 - \alpha$, kjer je $Z = \frac{\overline{X}-\mu}{\sigma}\sqrt{n}$. Tedaj je

$$P\left(\overline{X} - \frac{b\sigma}{\sqrt{n}} \le \mu \le \overline{X} - \frac{a\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Označimo $A = \overline{X} - \frac{b\sigma}{\sqrt{n}}$ in $B = \overline{X} - \frac{a\sigma}{\sqrt{n}}$.

Za katera a in b je interval $[A, B]$ najkrajši?

Pokazati je mogoče (Lagrangeova funkcija),
da mora biti $a = -b$ in $\Phi(b) = (1-\alpha)/2$;
oziroma, če označimo $b = z_{\alpha/2}$, velja $P(Z > z_{\alpha/2}) = \alpha/2$.

Iskani interval je torej
$A = \overline{X} - z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$ in $B = \overline{X} + z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$

tj., z verjetnostjo $1-\alpha$ je $|\overline{X} - \mu| < z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

Od tu dobimo, da mora za to,
da bo napaka manjša od ε z verjetnostjo $1-\alpha$,
veljati $n > (\frac{z_{\alpha/2}\sigma}{\varepsilon})^2$.

... Intervalsko ocenjevanje parametrov

Če pri porazdelitvi $X : N(\mu, \sigma)$ tudi parameter σ ni znan, ga nadomestimo s cenilko S in moramo zato uporabiti Studentovo statistiko $T = \frac{\overline{X}-\mu}{S}\sqrt{n}$. Ustrezni interval je tedaj

$$A = \overline{X} - t_{\alpha/2}\frac{S}{\sqrt{n}}, \qquad B = \overline{X} + t_{\alpha/2}\frac{S}{\sqrt{n}}$$

kjer je $P(T > t_{\alpha/2}) = \alpha/2$.

Če pa bi ocenjevali parameter σ^2, uporabimo statistiko $\chi^2 = (n-1)\frac{S^2}{\sigma^2}$, ki je porazdeljena po $\chi^2(n-1)$. Tedaj sta

$$A = \frac{(n-1)S^2}{b}, \qquad B = \frac{(n-1)S^2}{a}$$

Konstanti a in b včasih določimo iz pogojev $P(\chi^2 < a) = \alpha/2$ in $P(\chi^2 > b) = \alpha/2$; najkrajši interval pa dobimo, ko velja zveza $a^2 p(a) = b^2 p(b)$ in seveda $\int_a^b p(t)dt = 1 - \alpha$.

Teoretična interpretacija koeficienta zaupanja $(1-\alpha)$

Če zaporedoma izbiramo vzorce velikosti n
iz dane populacije in konstruiramo
$[(1-a)100]\%$ interval zaupanja
za vsak vzorec, potem lahko pričakujemo,
da bo $[(1-a)100]\%$ intervalov
dalo prvo vrednost parametra.

stopnja tveganja = 1 − stopnja zaupanja

I. $(1-\alpha)\%$-ni interval zaupanja za povprečje μ populacije, kadar poznamo standardni odklon σ:

točki

$$\overline{y} \pm z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

prestavljata krajišči intervala zaupanja, pri čemer je

$z_{\alpha/2}$ vrednost spremenljivke,
ki zavzame površino $\alpha/2$ na svoji desni;

σ je standardni odklon za populacijo;

n je velikost vzorca;

$\overline{y}$ je vrednost vzorčnega povprečja.

II. Veliki vzorec za $(1-\alpha)\%$-ni interval zaupanja za povprečje μ populacije:

$$\overline{y} \pm z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right),$$

kjer je s standardni odklon vzorca.

Primer: Na vzorcu velikosti $n = 151$ podjetnikov v majhnih podjetjih v Sloveniji, ki je bil izveden v okviru ankete 'Drobno gospodarstvo v Sloveniji' (Prašnikar, 1993), so izračunali, da je povprečna starost anketiranih podjetnikov $\overline{X} = 40,4$ let in standardni odklon $s = 10,2$ let. Pri 5 % tveganju želimo z intervalom zaupauja oceniti povprečno starost podjetnikov v majhnih podjetjih v Sloveniji.

$$P\left(\overline{y} - z_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \overline{y} + z_{\alpha/2}\frac{s}{\sqrt{n}}\right) = 1 - \alpha.$$

$$40,4 - \frac{1,96 \times 10,2}{\sqrt{151}} < \mu < 40,4 + \frac{1,96 \times 10,2}{\sqrt{151}}$$

$$40,4 - 1.6 < \mu < 40,4 + 1,6$$

95 % interval zaupanja za povprečno starost podjetnikov v majhnih podjetjih v Sloveniji je med $38,8$ in $42,0$ leti.

III. Majhen vzorec za $(1-\alpha)\%$-ni interval zaupanja za povprečje μ populacije:

$$\overline{y} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right),$$

kjer je porazdelitev spremenljivke y vzeta na osnovi $(n-1)$ prostostnih stopenj.

Privzeli smo: populacija, iz katere smo izbrali vzorec, ima **približno normalno porazdelitev**.

Primer: Vzemimo, da se spremenljivka X - število ur branja dnevnih časopisov na teden - porazdeljuje normalno $N(\mu, \sigma)$.

Na osnovi podatkov za 7 slučajno izbranih oseb ocenimo interval zaupanja za aritmetično sredino pri 10% tveganju.

Podatki in ustrezni izračuni so:

x_i	$x_i - \overline{X}$	$(x_i - \overline{X})^2$
5	-2	4
7	0	0
9	2	4
7	0	0
6	-1	1
10	3	9
5	-2	4
49	0	22

$$\overline{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{49}{7} = 7,$$

$$s^2 = \frac{\sum_{i=1}^n (x_i - \overline{X})^2}{n-1} = \frac{22}{6} = 3,67.$$

Iz tabele za t–porazdelitev preberemo, da je

$$t_{\alpha/2}(n-1) = t_{0,05}(6) = 1,943$$

in interval zaupanja je

$$7 - 1,943 \cdot \frac{1,9}{\sqrt{7}} < \mu < 7 + 1,943 \cdot \frac{1,9}{\sqrt{7}}$$

$$7 - 1,4 < \mu < 7 + 1,4.$$

IV. $(1-\alpha)$%-ni interval zaupanja za razliko $\mu_1 - \mu_2$, če poznamo odklona σ_1 in σ_2 in sta vzorca izbrana neodvisno:

$$\overline{y}_1 - \overline{y}_2 \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

V. Veliki vzorec za $(1-\alpha)$%-ni interval zaupanja za razliko $\mu_1 - \mu_2$, kadar ne poznamo odklonov σ_1 in σ_2, vzorce pa izbramo neodvisno:

$$\overline{y}_1 - \overline{y}_2 \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

VI. Majhen vzorec za $(1-\alpha)$%-ni interval zaupanja za razliko $\mu_1 - \mu_2$, kadar ne poznamo odklonov σ_1 in σ_2, ki pa sta si enaka, vzorci pa so majhni in izbrani neodvisno:

$$\overline{y}_1 - \overline{y}_2 \pm t_{\alpha/2, n_1+n_2-2}\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

kjer je $\displaystyle s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}.$

Privzeli smo:

- obe populaciji sta **približno normalni**,
- varianci sta enaki,
- naključni vzorci so izbrani **neodvisno**.

VII. Majhen vzorec za $(1-\alpha)\%$-ni interval zaupanja za razliko $\mu_1-\mu_2$, kadar ne poznamo odklonov σ_1 in σ_2, vzorci pa so majhni in izbrani neodvisno:

$$\overline{y}_1-\overline{y}_2\pm t_{\alpha/2,\nu}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

kjer je

$$\nu=\left\lfloor\frac{\left(s_1^2/n_1+s_2^2/n_2\right)^2}{(s_1^2/n_1)^2/(n_1-1)+(s_2^2/n_2)^2/(n_2-1)}\right\rfloor$$

Primer:

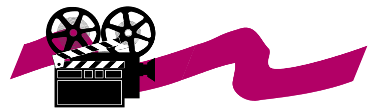
Naslednji podatki predstavljajo dolžine filmov, ki sta jih naredila dva filmska studija.

Izračunaj 90%-ni interval zaupanja za razliko med povprečnim časom filmov, ki sta jih producirala ta dva studija.

Predpostavimo, da so dolžine filmov porazdeljene **približno normalno**.

Čas (v minutah)

Studio 1:	103	94	110	87	98		
Studio 2:	97	82	123	92	175	88	118



Podatke vnesemo v Minitab

Film.MTV

```
Studio 1:   Studio 2:

 103           97
  94           82
 110          123
  87           92
  98          175
               88
              118
```

Dva vzorca T-Test in interval zaupanja

```
Dva vzorca T za C1 : C2
      N    povpr.   St.odk.   SE povpr.
C1    5    98.40     8.73       3.9
C2    7    110.7     32.2       12

90%-ni interval zaupanja za mu C1-mu C2:

T-TEST mu C1=mu C2 (vs ni=):
T = -0.96  P = 0.37  DF = 7
```

VIII. $(1-\alpha)\%$-ni interval zaupanja za razliko $\mu_d=\mu_1-\mu_2$ ujemajočih se parov v velikih vzorcih:

$$\overline{d}\pm z_{\alpha/2}\left(\frac{s_d}{\sqrt{n}}\right),$$ kjer je n število parov.

IX. $(1-\alpha)$%-ni interval zaupanja za razliko $\mu_d = \mu_1 - \mu_2$ ujemajočih se parov v majhnih vzorcih:

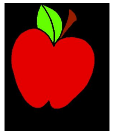
$$\overline{d} \pm t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right),\quad \text{kjer je } n \text{ število parov.}$$

Privzeli smo: populacija razlik parov je normalno porazdeljena.

Nal. 8-39. Špricanje jabolk lahko pozroči kontaminacijo zraka. Zato so v času najbolj intenzivnega špricanja zbrali in analizirali vzorce zraka za vsak od 11ih dni.

Raziskovalci želijo vedeti ali se povprečje ostankov škropiv (diazinon) razlikuje med dnevom in nočjo.

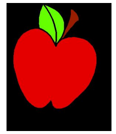
Analiziraj podatke za 90% interval zaupanja.



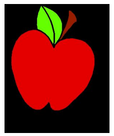
Nal. 8-39

Diazinon Residue

Datum	dan	no"c
Jan. 11	5,4	24,3
12	2,7	16,5
13	34,2	47,2
14	19,9	12,4
15	2,4	24,0
16	7,0	21,6
17	6,1	104,3
18	7,7	96,9
19	18,4	105,3
20	27,1	78,7
21	16,9	44,6

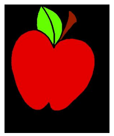


	Diazinon Residue		razlika
Datum	dan	no"c	dan-no"c
Jan. 11	5,4	24,3	-18,9
12	2,7	16,5	-13,8
13	34,2	47,2	-13,0
14	19,9	12,4	7,5
15	2,4	24,0	-21,6
16	7,0	21,6	-14,6
17	6,1	104,3	-98,2
18	7,7	96,9	-89,2
19	18,4	105,3	-86,9
20	27,1	78,7	-51,6
21	16,9	44,6	-27,7



Podatke vnesemo v Minitab:
Ex8-39.MTW

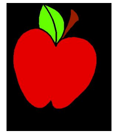
C1	C2
5.4	24.3
2.7	16.5
34.2	47.2
19.2	12.4
2.4	24.0
7.0	21.6
6.1	104.3
7.7	96.9
18.4	105.3
27.1	78.7
16.9	44.6




```
MTB > Let C3=C1-C2.

T interval zaupanja

Spremen. N   povpr. Stdev SEpovpr.
C3      11   -38.9   36.6    11.0
```



90,0 % interval zaupanja je $(58,9\,;\,18,9)$.

Za deleže

$p =$ delež populacije

$\hat{p} =$ delež vzorca,

kjer je $\hat{p} = y/n$ in je y število uspehov v n poskusih.

X. $(1-\alpha)\%$-ni interval zaupanja za delež populacije p, kadar poznamo $\sigma_{\hat{p}}$:

$$\hat{p} \pm z_{\alpha/2}\sigma_{\hat{p}}$$

XI. Veliki vzorec za $(1-\alpha)\%$-ni interval zaupanja za delež populacije:

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Privzeli smo: velikost vzorca n je dovolj velika, da je aproksimacija veljavna.

Dobro pravilo (angl. rule of thumb) za izpolnitev pogoja

“dovolj velik vzorec” je:

$$n\hat{p} \geq 4 \quad \text{in} \quad n\hat{q} \geq 4.$$

Primer: Na vzorcu ($n = 151$), ki je bil izveden v okviru ankete ’Drobno gospodarstvo v Sloveniji’, so izračunali. da je delež obrtnih podjetij $p = 0,50$. Pri 5 % tveganju želimo z intervalom zaupanja oceniti delež obrtnih majhnih podjetij v Sloveniji.

$$0,50 - 1.96\,\frac{0,50 \times 0,50}{\sqrt{151}} \;<\; \pi \;<\; 0,50 + 1.96\,\frac{0,50 \times 0,50}{\sqrt{151}}$$

$$0,50 - 0,08 \;<\; \pi \;<\; 0,50 + 0,08$$

S 5% stopnjo tveganja trdimo, da je delež obrtnih majhnih podjetij v Sloveniji glede na vsa majhna podjetja med 0,42 in 0,58.

XII. $(1-\alpha)\%$-ni interval zaupanja za razliko deležev $p_1 - p_2$, kadar poznamo $\sigma_{\hat{p}_1-\hat{p}_2}$:

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sigma_{\hat{p}_1-\hat{p}_2}.$$

XIII. Veliki vzorec za $(1-\alpha)\%$-ni interval zaupanja za razliko deležev $p_1 - p_2$:

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Privzeli smo: velikost vzorca n je dovolj velika, da je aproksimacija veljavna.

Kot splošno pravilo za dovolj velika vzorca privzamemo naslednje:

$$\begin{array}{ccc} n_1\hat{p}_1 \geq 4 & & n_1\hat{q}_1 \geq 4. \\ n_2\hat{p}_2 \geq 4 & \text{in} & n_2\hat{q}_2 \geq 4. \end{array}$$

XIV. Veliki vzorec za $(1-\alpha)\%$-ni interval zaupanja za varianco populacije σ^2:

$$\frac{(n-1)s^2}{\chi^2_{(\alpha/2,n-1)}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{(1-\alpha/2,n-1)}}$$

Privzeli smo: populacija iz katere izbiramo vzorce, ima **približno normalno porazdelitev**.

Primer: Vzemimo prejšnji primer spremenljivke o številu ur branja dnevnih časopisov na teden. Za omenjene podatke iz vzorca ocenimo z intervalom zaupanja varianco pri 10% tveganju.

Iz tabele za χ^2–porazdelitev preberemo, da je

$$\chi^2_{1-\alpha/2}(n-1) = \chi^2_{0,95}(6) = 12,6,$$

$$\chi^2_{\alpha/2}(n-1) = \chi^2_{0,05}(6) = 1,64.$$

90 % interval zaupanja za varianco je tedaj

$$\frac{6 \cdot 3,67}{12,6} < \sigma^2 < \frac{6 \cdot 3,67}{1,64},$$

$$1,75 < \sigma^2 < 13,43.$$

XV. $(1-\alpha)\%$-ni interval zaupanja za kvocijent varianc dveh populacij σ_1^2/σ_2^2:

$$\frac{s_1^2}{s_2^2} \cdot \frac{1}{F_{(\alpha/2,n_1-1,n_2-1}} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2} \cdot \frac{1}{F_{(\alpha/2,n_2-1,n_1-1}}$$

Privzeli smo:

– obe populaciji iz katerih izbiramo vzorce, imata **približno normalni porazdelitvi** relativnih frekvenc.

– naključni vzorci so izbrani **neodvisno** iz obeh populacij.

XVI. Izbira velikosti vzorca za oceno populacijskega povprečja μ znotraj H enot z verjetnostjo $(1-\alpha)$:

$$n = \left(\frac{z_{\alpha/2}\sigma}{H}\right)^2$$

Populacijski odklon mora biti običajno aproksimiran.

XVII. Izbira velikosti vzorca za oceno razlike $\mu_1 - \mu_2$ med parom populacijskih povprečij, ki je pravilna znotraj H enot z verjetnostjo $(1-\alpha)$:

$$n_1 = n_2 = \left(\frac{z_{\alpha/2}}{H}\right)^2 (\sigma_1^2 + \sigma_2^2)$$

XVIII. Izbira velikosti vzorca za oceno deleža populacije p, ki je pravilna znotraj H enot z verjetnostjo $(1-\alpha)$:

$$n = \left(\frac{z_{\alpha/2}}{H}\right)^2 pq$$

Opozorilo: v tem primeru potrebujemo oceni za p in q.
Če nimamo nobene na voljo, potem uporabimo $p = q = 0,5$ za konzervativno izbiro števila n.

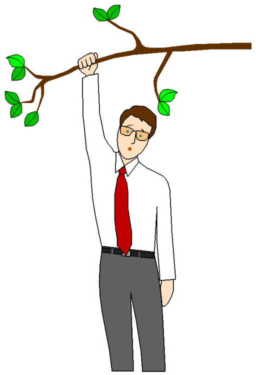
XIX. Izbira velikosti vzorca za cenilko razlike $p_1 - p_2$ med dvema deležema populacije, ki je pravilna znotraj H enot z verjetnostjo $(1-\alpha)$:

$$n_1 = n_2 = \left(\frac{z_{\alpha/2}}{H}\right)^2 (p_1 q_1 + p_2 q_2)$$

II.5. Preizkušanje statističnih domnev

Načrt

- postopek
- elementi
  - napake 1. in 2. vrste
  - značilno razlikovanje
  - moč statističnega testa
- testi
  - centralna tendenca
  - delež
  - varianca



Uvod

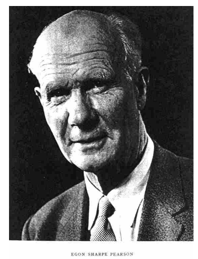
- postavimo trditev o populaciji,
- izberemo vzorec,
  s katerim bomo preverili trditev,
- zavrni ali sprejmi trditev.

Hipoteza je testirana z določanjem verjetja, da dobimo določen rezultat, kadar jemljemo vzorce iz populacije s predpostavljenimi vrednostimi parametrov.

Zgodovina

Teorijo preizkušanja domnev sta v 20. in 30. letih prejšnjega stoletja razvila J. Neyman in E.S. Pearson.

Statistična domneva (ali hipoteza) je vsaka domneva o porazdelitvi slučajne spremenljivke X na populaciji.



Če poznamo vrsto (obliko) porazdelitve $f(x;\zeta)$ in postavljamo/raziskujemo domnevo o parametru ζ, govorimo o *parametrični domnevi*.

Če pa je vprašljiva tudi sama vrsta porazdelitve,
je domneva *neparametrična*.

Domneva je *enostavna*, če natačno določa porazdelitev
(njeno vrsto in točno vrednost parametra); sicer je *sestavljena*.

Primer: Naj bo $X : N(\mu, \sigma)$.

Če poznamo σ, je domneva $H : \mu = 0$ enostavna;

Če pa parametra σ ne poznamo, je sestavljena.

Primer sestavljene domneve je tudi $H : \mu > 0$.

Statistična domneva je lahko pravilna ali napačna.
Želimo seveda sprejeti pravilno domnevo in zavrniti napačno.
Težava je v tem, da o pravilnosti/napačnosti domneve
ne moremo biti gotovi, če jo ne preverimo na celotni populaciji.
Ponavadi se odločamo le na podlagi vzorca.

Če vzorčni podatki preveč odstopajo od domneve, rečemo,
da niso *skladni* z domnevo, oziroma,
da so *razlike značilne*, in domnevo zavrnemo.

Če pa podatki domnevo podpirajo, jo ne zavrnemo
– včasih jo celo sprejmemo.
To ne pomeni, da je domneva pravilna, temveč
da ni zadostnega razloga za zavrnitev.

Postopek testiranja hipoteze

- postavi ničelno in alternativno hipotezo,
- izberi testno statistiko,
- določi zavrnitveni kriterij,
- izberi naključni vzorec,
- izračunaj vrednost na osnovi testne statistike,
- sprejmi odločitev,
- naredi ustrezen zaključek.

Hipoteza

- Ničelna hipoteza (H_0)
  - je trditev o lastnosti populacije za katero predpostavimo, da drži (oziroma za katero verjamemo, da je resnična),
  - je trditev, ki jo test skuša ovreči.
- Alternativna (nasprotna) hipoteza (H_a)
  - je trditev nasprotna ničelni hipotezi,
  - je trditev, ki jo s testiranjem skušamo dokazati.

... Hipoteza

- ničelna hipoteza (H_0)
  - obtoženec je nedolžen,
- alternativna hipoteza (H_a)
  - obtoženec je kriv.

Odločitev in zaključek

- Porota je spoznala obtoženca za **krivega**.
  Zaključimo, da je bilo dovolj dokazov, ki nas prepričajo,
  da je obtoženec storil kaznivo dejanje.
- Porota je spoznala obtoženca za **nedolžnega**.
  Zaključimo, da je ni bilo dovolj dokazov, ki bi nas prepričali,
  da je obtoženec storil kaznivo dejanje.

Elementi testiranja hipoteze

		odločitev	
		nedolžen	**kriv**
dejansko stanje	**nedolžen**	pravilna odločitev	napaka 1. vrste (α)
	kriv	napaka 2. vrste (β)	moč ($1-\beta$)

... Elementi testiranja hipoteze

- verjetnost napake 1. vrste (α)
  verjetnost za obtožbo nedolžnega obtoženca.
- značilno razlikovanje (signifikantno) oziroma **stopnja značilnosti**
- količina dvoma (α), ki ga bo porota še sprejela.
  - Kriminalna tožba: Beyond a reasonable doubt...
  - Civilna tožba: The preponderance of evidence must suggest...

... Elementi testiranja hipoteze

- verjetnost napake 2. vrste: (β)
  - verjetnost, da spoznamo krivega obtoženca za nedolžnega,
- moč testa: ($1-\beta$)
  - verjetnost, da obtožimo krivega obtoženca.

Sodba

- breme dokazov,
- potrebno je prepričati poroto, da je obtoženi kriv
  (alternativna hipoteza)
  preko določene stopnje značilnosti.
  - Criminal: Reasonable Doubt
  - Civil: Preponderance of evidence

Obramba

- Ni bremena dokazovanja.
- Povzročiti morajo dovolj dvoma pri poroti,
  če je obtoženi resnično kriv.

Statistična ničelna hipoteza

$H_0 \; : \; \mu = 9mm$
(Premer 9 milimetrskega kroga),

$H_0 \; : \; \mu =$ 600 km
(Proizvalajec trdi, da je to doseg novih vozil),

$H_0 \; : \; \mu =$ 3 dnevi

Neusmerjena alternativna hipoteza

$H_0 \ : \ \mu = 9mm$

$H_a \ : \ \mu \neq 9mm$

Premer 9 milimetrskega kroga.

“Manj kot” alternativna hipoteza

$H_0 \ : \ \mu = 600$ km

$H_a \ : \ \mu < 600$ km

Proizvalajec trdi, da je to
doseg novih vozil.

“Več kot” alternativna hipoteza

$H_0 \ : \ \mu = 3$ dnevi

$H_a \ : \ \mu > 3$ dnevi

Čas odsotnosti določenega artikla
pri neposredni podpori.

Definicije

1. Zavrnitev ničelne hipoteze, če je le-ta pravilna, je **napaka 1. vrste.**
   Verjetnost, da naredimo napako 1. vrste, označimo s simbolom α in ji pravimo **stopnja tveganja**, $(1 - \alpha)$ pa je **stopnja zaupanja**.
2. Če ne zavrnemo ničelno hipotezo, v primeru, da je napačna, pravimo, da gre za **napako 2. vrste**.
   Verjetnost, da naredimo napako 2. vrste, označimo s simbolom β.
3. **Moč statističnega testa**, $(1 - \beta)$
   je verjetnost zavrnitve ničelne hipoteze v primeru, ko je le-ta v resnici napačna.

Statistična hipoteza

- ničelna hipoteza
  - $H_0 \ : \ q = q_0$
- alternativna hipoteza
  - $H_a \ : \ q \neq q_0$
  - $H_a \ : \ q > q_0$
  - $H_a \ : \ q < q_0$

Primer testiranja hipoteze

Predpostavimo, da je dejanska mediana (τ) pH iz določene regije 6,0.

Da bi preverili to trditev, bomo izbrali 10 vzorcev zemlje iz te regije, da ugotovimo, če empirični vzorci močno podpirajo, da je dejanska mediana manšja ali enaka 6,0?

Predpostavke

- naključni vzorec
  - neodvisen
  - enako porazdeljen (kot celotna populacija),
- vzorčenje iz zvezne porazdelitve,
- verjetnostna porazdelitev ima mediano.

Postavitev statistične hipoteze in izbira testne statistike

- ničelna hipoteza
  - $H_0 \,:\, \tau = 6,0$ (mediana populacije τ_0)
- alternativna hipoteza
  - $H_a \,:\, \tau < 6,0$

Testna statistika (TS)

- $S_+ =$ število vzorcev, ki so večji od mediane τ_0 iz hipoteze,
- $S_- =$ število vzorcev, ki so manjši od mediane τ_0 iz hipoteze.

Porazdelitev testne statistike

- vsak poskus je bodisi uspeh ali neuspeh,
- fiksen vzorec, velikosti n,
- naključni vzorci
  - neodvisni poskusi,
  - konstantna verjetnost uspeha.

Torej gre za

- binomsko porazdelitev: $S_+ \approx B(n,p)$,
- s parameteri $n = 10$ in $p = 0,5$,
- in pričakovano vrednostjo (matematičnim upanjem): $E(X) = np = 5$.

Testiranje hipoteze

		odločitev	
		FTR H_0	**zavrni** H_0
dejansko stanje	H_0 **je pravilna**	pravilna odločitev	
	H_0 **je napačna**		pravilna odločitev

Napaka 1. vrste

		odločitev	
		FTR H_0	**zavrni** H_0
dejansko stanje	H_0 **je pravilna**		napaka 1. vrste
	H_0 **je napačna**		

Napaka 2. vrste

		odločitev	
		FTR H_0	**zavrni** H_0
dejansko stanje	H_0 **je pravilna**		
	H_0 **je napačna**	napaka 2. vrste	

Moč testa

		odločitev	
		FTR H_0	**zavrni** H_0
dejansko stanje	H_0 **je pravilna**		
	H_0 **je napačna**		$(1-\beta)$

Elementi testiranja hipoteze

		odločitev	
		FTR H_0	**zavrni** H_0
dejansko stanje	H_0 **je pravilna**		(α)
	H_0 **je napačna**	(β)	$(1-\beta)$

... Elementi testiranja hipoteze

- verjetnost napake 1. vrste (α)
  - Če hipoteza H_0 drži, kakšna je možnost, da jo zavržemo.
- stopnja značilnosti testa (signifikantnosti)
  - Največji α, ki ga je vodja eksperimenta pripravljen sprejeti (zgornja meja za napako 1. vrste).
- verjetnost napake 2. vrste (β)
  - Če hipoteza H_0 ne drži, kakšna je možnost, da je **ne** zavržemo.
- moč statističnega testa: $(1-\beta)$
  - Če hipoteza H_0 ne drži, kakšna je možnost, da jo zavržemo.

... Elementi testiranja hipoteze

velikost vzorca	napaka 1.vrste	napaka 2.vrste	moč
n	α	β	$1-\beta$
konst.	$\uparrow$	$\downarrow$	$\uparrow$
konst.	$\downarrow$	$\uparrow$	$\downarrow$
povečanje	$\downarrow$	$\downarrow$	$\uparrow$
zamnjšanje	$\uparrow$	$\uparrow$	$\downarrow$

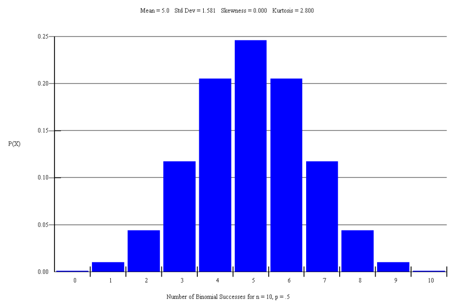
Primer (A) testiranja hipoteze

Predpostavimo, da je dejanska mediana (τ) pH iz določene regije 6,0.

Da bi preverili to trditev, bomo izbrali 10 vzorcev zemlje iz te regije, da ugotovimo, če empirični vzorci močno podpirajo, da je dejanska mediana manjša ali enaka 6,0?

- Hipotezi
  - $H_0 : \tau = 6,0$
  - $H_a : \tau < 6,0$
- Testna statistika
  - $S_+ =$ število vzorcev večjih od predpostavljene mediane τ_0,
  - $S_+ \approx B(n,p) = B(10; 0,5)$,
  - $E(S_+) = 5$.

Porazdelitev testne statistike



Določimo zavrnitveni kriterij

x	$P(X=x)$	$F(x)$
0	$0,000977$	$0,00098$
1	$0,009766$	$0,01074$
2	$0,043945$	$0,05469$
3	$0,117188$	$0,17188$
4	$0,205078$	$0,37695$
5	$0,246094$	$0,62305$
6	$0,205078$	$0,82813$
7	$0,117188$	$0,94531$
8	$0,043945$	$0,98926$
9	$0,009766$	$0,99902$
10	$0,000977$	$1,00000$

... Določimo zavrnitveni kriterij

- Stopnja značilnosti testa (α) = 0,01074,
- Kritična vrednost
  - $S_+ = 1$,
- Območje zavrnitve
  - $S_+ \leq 1$,

Izberemo naključni vzorec

Predpostavimo, da je dejanska mediana (τ) pH iz določene regije 6,0.

Da bi preverili to trditev, smo izbrali 10 vzorcev zemlje iz te regije in jih podvrgli kemični analizi in na ta način določili pH vrednost za vsak vzorec.

Ali empirični podatki podpirajo trditev,
da je dejanska mediana manšja ali enaka 6,0?

$$5,93;\ 6,08;\ 5,86;\ 5,91;\ 6,12;\ 5,90;\ 5,95;\ 5,89;\ 5,98;\ 5,96.$$

Izračunaj vrednost iz testne statistike

pH	predznak
5,93	−
6,08	+
5,86	−
5,91	−
6,12	+
5,90	−
5,95	−
5,89	−
5,98	−
5,96	−

$S_+ = 2$

Naredimo odločitev

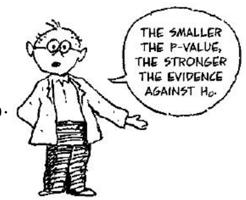
- Izračunana vrednost S_+ leži zunaj zavrnitvenega območja.
- Ni osnove za zavrnitev hipoteze H_0.

P-vrednost

P-vrednost (ali ugotovljena bistvena stopnja za določen statistični test) je verjetnost (ob predpostavki, da drži H_0), da ugotovimo vrednost testne statistike, ki je vsaj toliko v protislovju s H_0 in podpira H_a kot tisto, ki je izračunana iz vzorčnih podatkov.

P-vrednost

- Sprejemljivost hipoteze H_0 na osnovi vzorca
  - Verjetnost, da je opazovani vzorec (ali podatki) bolj ekstremni, če je hipoteza H_0 pravilna.
- Najmanjši α pri katerem zavrnemo hipotezo H_0.
  - če je P-vrednost $> \alpha$, potem FTR H_0,
  - če je P-vrednost $< \alpha$, potem zavrni H_0.



... P-vrednost

x	$P(X=x)$	$F(x)$
0	$0,000977$	$0,00098$
1	$0,009766$	$0,01074$
2	$0,043945$	$0,05469$
3	$0,117188$	$0,17188$
4	$0,205078$	$0,37695$
5	$0,246094$	$0,62305$
6	$0,205078$	$0,82813$
7	$0,117188$	$0,94531$
8	$0,043945$	$0,98926$
9	$0,009766$	$0,99902$
10	$0,000977$	$1,00000$

P-vrednost $= P(S_+ \geq 2 \,|\, \tau = 6,0) = 0,05469$.

Izračunaj vrednost iz testne statistike

pH	predznak
$5,93$	$-$
$6,08$	$+$
$5,86$	$-$
$5,91$	$-$
$6,12$	$+$
$5,90$	$-$
$5,95$	$-$
$5,89$	$-$
$5,98$	$-$
$5,96$	$-$

$S_- = 8$

... P-vrednost

x	$P(X=x)$	$F(x)$
0	$0,000977$	$0,00098$
1	$0,009766$	$0,01074$
2	$0,043945$	$0,05469$
3	$0,117188$	$0,17188$
4	$0,205078$	$0,37695$
5	$0,246094$	$0,62305$
6	$0,205078$	$0,82813$
7	$0,117188$	$0,94531$
8	$0,043945$	$0,98926$
9	$0,009766$	$0,99902$
10	$0,000977$	$1,00000$

P-vrednost $= P(S_- \geq 8 \,|\, \tau = 6,0) = 0,05469$.

Odločitev in zaključek

- P-vrednost $> \alpha = 0,01074$.
- Ni osnove za zavrnitev hipoteze H_0.
- Zavrni ničelno hipotezo.
  - Zaključimo, da empirični podatki sugerirajo, da velja alternativna trditev.
- Ni osnove za zavrnitev ničelne hipoteze (angl. fail to reject, kratica FTR).
  - Zaključimo, da nimamo dovolj osnov, da bi dokazali, da velja alternativna trditev.
- Premalo podatkov, da bi pokazali, da je dejanska mediana pH manjša od 6,0.
- Privzemimo, da je pH enaka 6,0 v tej konkretni regiji.

Naloga 9.4 na strani 429

Pascal je visoko-nivojski programski jezik, ki smo ga nekoč pogosto uporabljali na miniračunalnikih in microprocesorjih.

Narejen je bil eksperiment, da bi ugotovili delež Pascalovih spremenljivk, ki so tabelarične spremenljivke (v kontrast skalarim spremenljivkam, ki so manj učinkovite, glede na čas izvajanja).

20 spremenljivk je bilo naključno izbranih iz množice Pascalovih programov, pri tem pa je bilo zabeleženo število tabelaričnih spremenljivk Y.

Predpostavimo, da želimo testirati hipotezo, da je Pascal bolj učinkovit jezik kot Agol, pri katerem je 20% spremenljivk tabelaričnih.

To pomeni, da bomo testirali $H_0 \,:\, p = 0,20$, proti $H_a \,:\, p > 0,20$, kjer je p verjetnost, da imamo tabelarično spremenljivko na vsakem poskusu.

Predpostavimo, da je 20 poskusov neodvisnih.

(a) Določi α za območje zavrnitve $y > 8$.

Izračunati želimo verjetnost, da se bo zgodila napaka 1. vrste, torej da bomo zavrnili pravilno hipotezo.

Predpostavimo, da je hipoteza H_0 pravilna, tj. $Y\,:\,\mathrm{B}(20; 0{,}2)$. Če se bo zgodilo, da bo Y pri izbranem vzorcu večji ali enak 8, bom hipotezo zavrnili, čeprav je pravilna. Torej velja:

$$\begin{aligned}\alpha &= \mathrm{P}(Y \geq 8) = 1 - \mathrm{P}(Y \leq 7)\\ &= 1 - \sum_{i=0}^{7} \mathrm{P}(Y = i)\\ &= 1 - \sum_{i=0}^{7} \binom{20}{i}\, 0{,}2^i\, 0{,}2^{20-i}\\ &= 1 - 0{,}9679 = 0{,}0321 = 3{,}21\%.\end{aligned}$$

(b) Določi α za območje zavrnitve $y \geq 5$.

Do rezultata pridemo na enak način kot v prejšnji točki:

$$\begin{aligned}\alpha &= \mathrm{P}(Y \geq 5) = 1 - \mathrm{P}(Y \leq 4)\\ &= 1 - \sum_{i=0}^{4} \mathrm{P}(Y = i)\\ &= 1 - \sum_{i=0}^{4} \binom{20}{i} 0{,}2^i\, 0{,}2^{20-i}\\ &= 1 - 0{,}6296 = 0{,}3704 = 37{,}04\%.\end{aligned}$$

(c) Določi β za območje zavrnitve $Y \geq 8$, če je $p = 0,5$.

Izračunati želimo verjetnost, da se bo zgodila napaka 2. vrste, torej da bomo sprejeli napačno hipotezo. Ker vemo, da je $p = 0{,}5$, velja $Y \sim \mathrm{B}(20; 0{,}5)$.

Napačno hipotezo bomo sprejeli, če bo y pri izbranem vzorcu manjši od 8.

$$\beta = \mathrm{P}(y \leq 7) = \sum_{i=0}^{7} \binom{20}{i} 0{,}5^i\, 0{,}5^i = 0{,}1316 = 13{,}16\%.$$

(d) Določi β za območje zavrnitve $y \geq 5$, če je $p = 0{,}5$.

Do rezultata pridemo na enak način kot v prejšnji točki:

$$\beta = \mathrm{P}(y \leq 4) = \sum_{i=0}^{4} \binom{20}{i} 0{,}5^i\, 0{,}5^i = 0{,}0059 = 0{,}59\%.$$

(e) Katero območje zavrnitve $y \geq 8$ ali $y \geq 5$ je bolj zaželjeno, če želimo minimizirati verjetnost napake 1. stopnje oziroma če želimo minimizirati verjetnost napake 2. stopnje.

Napako 1. stopnje minimiziramo z izbiro območja $y \geq 8$, napako 2. stopnje pa z izbiro območja $y \geq 5$.

(f) Določi območje zavrnitve $y \geq a$ tako, da je α približno 0,01.

Na osnovi točke (e) zaključimo, da se z večanjem števila a manjša verjetnost α in s poskušanjem (ki ga pričnemo na osnovi izkušenje iz točke (a) pri 9) pridemo do $a = 9$.

(g) Za območje zavrnitve določeno v točki (f) določi moč testa,
če je v resnici $p = 0,4$.

Moč testa je $1-\beta$. Verjetnost β izračunamo enako kot v točkah (c) in (d). Velja $Y \sim \mathrm{B}(20; 0{,}4)$ in

$$\beta = \mathrm{P}(y \le 8) = \sum_{i=0}^{8} \binom{20}{i} 0{,}4^i \, 0{,}6^i = 0{,}5956 = 59{,}56\%.$$

Moč testa znaša $0{,}4044$.

(h) Za območje zavrnitve določeno v točki (f) določi moč testa, če je v resnici $p = 0{,}7$.

Tokrat velja $Y \sim \mathrm{B}(20; 0{,}4)$ in

$$\beta = \mathrm{P}(y \le 8) = \sum_{i=0}^{8} \binom{20}{i} 0{,}7^i \, 0{,}3^i = 0{,}0051 = 0{,}51\%.$$

Moč testa znaša $0{,}995$.

Formalen postopek za testiranje hipotez

1. Postavi hipotezi:
   - ničelna,
   - alternativna.
2. Za parameter poiščemo kar se da dobro cenilko (npr. nepristransko) in njeno porazdelitev ali porazdelitev ustrezne statistike (izraz, v katerem nastopa cenilka).
3. Določi odločitveno pravilo.
   Izberemo stopnjo značilnosti (α).
   Na osnovi stopnje značilnosti in porazdelitve statistike določimo kritično območje;
4. Zberi/manipuliraj podatke ter na vzorčnih podatkih izračunaj (eksperimentalno) vrednost testne statistike.

5. Primerjaj in naredi zaključek.

- če eksperimentalna vrednost pade v kritično območje, ničelno domnevo zavrni in sprejmi osnovno domnevo ob stopnji značilnosti α.
- če eksperimentalna vrednost ne pade v kritično območje, pa pravimo da vzorčni podatki kažejo na statistično neznačilne razlike med parametrom in vzorčno oceno.

I. $H_0 \; : \; \mu = \mu_0$

Če poznamo odklon σ, potem

$$\mathrm{T.S.} \; = \frac{\bar{y} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$ **sledi z-porazdelitev.**

Primer (B)

Proizvajalec omake za špagete da v vsako posodo 28 unče omake za špagete. Količina omake, ki je v vsaki posodi, je porazdeljena normalno s standardnim odklonom $0,005$ unče.

Podjetje ustavi proizvodni trak in popravi napravo za polnenje, če so posode bodisi

– premalo napolnjene (to razjezi kupce),

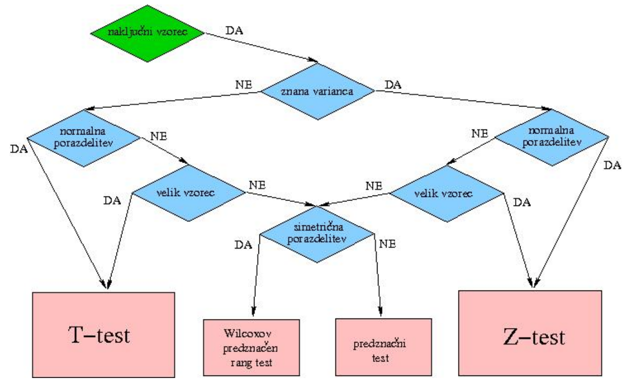
– ali preveč napolnjene (kar seveda pomeni manjši profit).

Ali naj na osnovi vzorca iz 15ih posod ustavijo proizvodno linijo? Uporabi stopnjo značilnosti $0,05$.

Postavimo hipotezo

- ničelna hipoteza
  - $H_0 \; : \; \mu = 28$
- alternativna hipoteza
  - $H_a \; : \; \mu \neq 28$

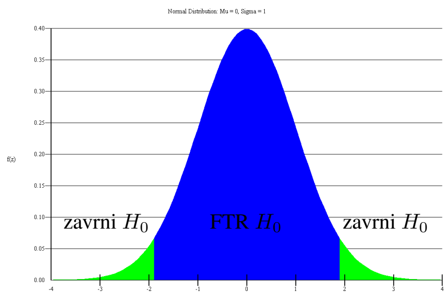
Izberimo testno statistiko



Z-Test

- test
  - $H_0 \; : \; \mu = \mu_0$ (povprečje populacije)
- predpostavke
  - naključno vzorčenje
  - poznamo varianco populacije
  - izbiramo vzorce iz normalne porazdelitve in/ali imamo vzorec pri katerem je n velik.

Določimo zavrnitveni kriterij



Rezultati testiranja

- naredi naključni vzorec
  - vzorčno povprečje: $28,0165$
- izračunaj vrednost testne statistike

$$Z = (28,0165 - 28)/0,0129 = 1,278$$

- naredi odločitev
  - FTR H_0
- zaključek
  - privzemi $\mu = 28$

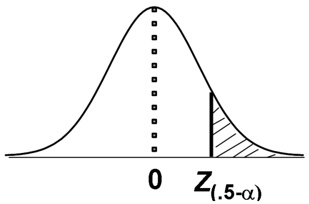
P-vrednost

- Sprejemljivost hipoteze H_0 na osnovi vzorca
  - možnost za opazovanje vzorca (ali bolj ekstremno podatkov), če je hipoteza H_0 pravilna
  - P-vrednost $= (2)P(Z > 1,278) = (2)(0,1003) = 0,2006$
- Najmanjši α pri katerem zavrnemo hipotezo H_0
  - P-vrednost $> \alpha$, zato FTR H_0

Za $H_a\ :\ \mu > \mu_0$

odločitveno pravilo: zavrni H_0, če je

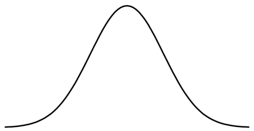
$$\text{T.S.}\ \geq z_{(0{,}5-\alpha)}$$



Za $H_a\ :\ \mu < \mu_0$

odločitveno pravilo: zavrni H_0, če je

$$\text{T.S.}\ \leq z_{(0{,}5-\alpha)}$$

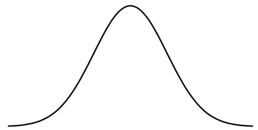


Za $H_a\ :\ \mu \neq \mu_0$

odločitveno pravilo: zavrni H_0

če je $\text{T.S.}\ \leq -z_{(0{,}5-\alpha)}$

ali če je $\text{T.S.}\ \geq z_{(0{,}5-\alpha)}$



II. $H_0\ :\ \mu = \mu_0$

Če ne poznamo odklona σ in je $n \geq 30$, potem

$$\text{T.S.}\ = \frac{\overline{y} - \mu_0}{\frac{s}{\sqrt{n}}} \quad \textbf{sledi } z\textbf{-porazdelitev.}$$

III. $H_0\ :\ \mu = \mu_0$

Če ne poznamo odklona σ, populacija je normalna in je $n < 30$, potem

$$\text{T.S.}\ = \frac{\overline{y} - \mu_0}{\frac{s}{\sqrt{n}}} \quad \textbf{sledi}$$

t-porazdelitev z $n-1$ prostostnimi stopnjami.

Primer (C2)

Za slučajni vzorec: 16-ih odraslih Slovencev smo izračunali povprečno število in variance priznanih let šolanja: $\overline{X} = 9$ in $s^2 = 9$. Predpostavljamo, da se spremenljivka na populaciji porazdeljuje normalno.

Ali lahko sprejmemo domnevo, da imajo odrasli Slovenci v povprečju več kot osemletko pri 5% stopnji značilnosti?

Postavimo najprej ničelno in osnovno domnevo:

$$H_0\ :\ \mu = 8 \quad \text{in} \quad H_1\ :\ \mu > 8.$$

Ustrezna statistike je

$$t = \frac{\overline{X} - \mu_H}{s}\sqrt{n},$$

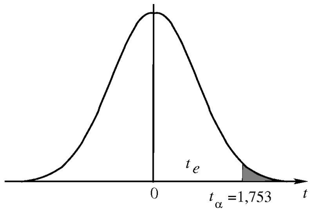
ki se porazdeljnje po t-porazdelitvi s 15 prostostnimi stopnjami.

Ker gre za enostranski test,
je glede na osnovno domnevo
kritično območje na desni strani
porazdelitve in kritična vrednost

$$t_{0,05}(15) = 1,753.$$

Izračunajmo eksperimentalno vrednost
statistike:

$$t_e = \frac{9-8}{3}\sqrt{16} = 1,3$$



Eksperimentalna vrednost ne pade v kritično območje.
Zato ničelne domneve ne moremo zavrniti
in sprejeti osnovne domneve,
da imajo odrasli Slovenci več kot osemletko.

Primer (C)

Ravnatelj bežigrajske gimnazije trdi, da imajo najboljši PT program v Sloveniji s povprečjem APFT 240.

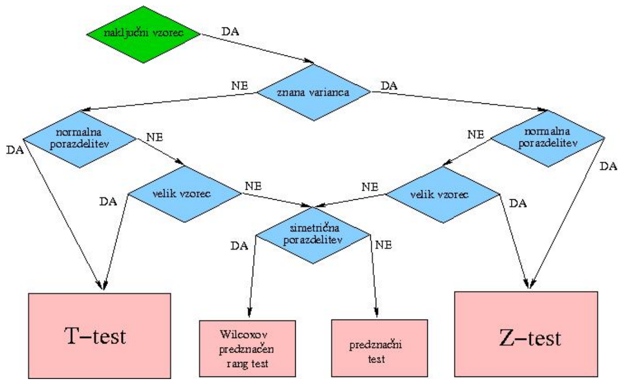
Predpostavi, da je porazdelitev rezultatov testov približno normalna.

Uporabi $\alpha = 0,05$ za določitev ali je povprečje APFT rezultatov šestih naključno izbranih dijakov iz bežigrajske gimnazije statistično večje od 240?

Postavimo hipotezo

- ničelna hipoteza
  - $H_0 : \mu = 240$
- alternativna hipoteza
  - $H_a : \mu > 240$

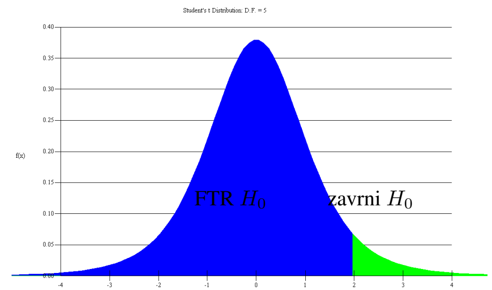
Izberimo testno statistiko



T-test

- test
  - $H_0 : \mu = \mu_0$ (povprečje populacije)
- predpostavke
  - naključno vzorčenje
  - ne poznamo varianco populacije
  - izbiramo vzorce iz normalne porazdelitve in/ali imamo vzorec pri katerem je n velik.

Določimo zavrnitveni kriterij



Rezultati testov

- naredi naključni vzorec
  - vzorčno povprečje: 255,4
  - vzorčni standardni odklon: 40,07
- izračunaj vrednost testne statistike

$$T = (255,4 - 240)/16,36 = 0,9413$$

- sprejmi odločitev
  - FTR H_0
- zaključek
  - Bežigrajska gimnazija ne more pokazati, da imajo višje povprečje APFT rezultatov, kot slovensko povprečje.

P-vrednost

- Sprejemljivost hipoteze H_0 na osnovi vzorca
  - možnost za opazovanje vzorca (ali bolj ekstremno podatkov), če je hipoteza H_0 pravilna
  - P-vrednost $= P(T > 0,9413) = 0,1949$.
- Najmanjši α pri katerem zavrnemo hipotezo H_0
  - P-vrednost $> \alpha$, zato FTR H_0 .

Vstavimo podatke v Minitab
(Ex9-23.MTV)

C1:
2610
2750
2420
2510
2540
2490
2680

T-test povprečja

```
Test of mu = 2500.0 vs mu > 2500.0

   N      MEAN    STDEV     SE MEAN
C1 7     2571.4   115.1        43.5

     T      p-VALUE
  1.64       0.076
```

Razlaga P-vrednosti

1. Izberi največjo vrednost za α, ki smo jo pripravljeni tolerirati.
2. Če je P-vrednost testa manjša kot maksimalna vrednost parametra α, potem zavrni ničelno hipotezo.

Razlika povprečij dveh populaciji

IV. $H_0 : \mu_1 - \mu_2 = D_0$

Če poznamo σ_1 in σ_2
ter jemljemo vzorce neodvisno, potem

$$\text{T.S.} = \frac{(\overline{y}_1 - \overline{y}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \textbf{sledi } z\textbf{-porazdelitev.}$$

Primer 1: Preveriti želimo domnevo, da so dekleta na izpitu boljša od fantov. To domnevo preverimo tako, da izberemo slučajni vzorec 36 deklet in slučajni vzorec 36 fantov, za katere imamo izpitne rezultate, na katerih izračunamo naslednje statistične karakteristike:

$$\overline{X}_F = 7,0, \quad s_F = 1$$

$$\overline{X}_D = 7,2, \quad s_D = 1$$

Domnevo preverimo pri 5% stopnji značilnosti.

Postavimo ničelno in osnovno domnevo:

$$H_0 : \mu_D = \mu_F \quad \text{oziroma} \quad \mu_D - \mu_F = 0,$$

$$H_1 : \mu_D > \mu_F \quad \text{oziroma} \quad \mu_D - \mu_F > 0.$$

Za popularijsko razliko aritmetičnih sredin na vzorcih računamo vzorčno razliko aritmetičnih sredin, ki se za dovolj velike vzorce porazdeljuje normalno

$$\overline{X}_D - \overline{X}_F : N\left(\mu_D - \mu_F, \sqrt{\frac{s_D^2}{n_D} + \frac{s_F^2}{n_F}}\right).$$

oziroma statistika

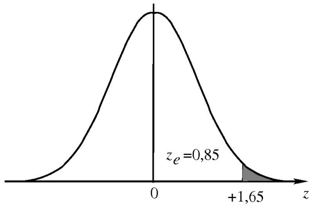
$$z = \frac{\overline{X}_D - \overline{X}_F - (\mu_D - \mu_F)_H}{\sqrt{\frac{s_D^2}{n_D} + \frac{s_F^2}{n_F}}}$$

standardizirano normalno $N(0, 1)$.

Osnovna domneva kaže enostranski test: možnost napake 1. vrste je le na desni strani normalne porazdelitve, kjer zavračamo ničelno domnevo. Zato je kritično območje določeno z vrednostmi večjimi od 1,65.

Eksperimentalna vrednost statistike je

$$z_e = \frac{7,2 - 7 - 0}{\sqrt{\frac{1}{36} + \frac{1}{36}}} = 0,852.$$



Eksperimentalna vrednost ne pade v kritično območje.
Ničelne domneve ne moremo zavrniti.
Povprečna uspešnost deklet in fantov ni statistično značilno različna.

V. $H_0 : \mu_1 - \mu_2 = D_0$

Če ne poznamo σ_1 in/ali σ_2,
vzorce jemljemo neodvisno,
$n_1 \geq 30$ in/ali $n_2 \geq 30$, potem

$$\text{T.S.} = \frac{(\overline{y}_1 - \overline{y}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad \textbf{sledi } z\textbf{-porazdelitev.}$$

VI. $H_0 : \mu_1 - \mu_2 = D_0$

Če ne poznamo σ_1 in/ali σ_2,
vzorce jemljemo neodvisno,
populaciji sta normalno porazdeljeni,
varianci obeh populacij sta enaki,
$n_1 < 30$ ali $n_2 < 30$, potem

T.S. =

$$\frac{(\overline{y}_1-\overline{y}_2)-D_0}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

sledi t-porazdelitev z n_1+n_2-2 prostostnimi stopnjami.

Privzeli smo:

1. Populaciji iz katerih jemljemo vzorce imata obe približno **normalno** relativno porazdelitev frekvenc.
2. Varianci obeh populacij sta **enaki**.
3. Naključni vzorci so izbrani **neodvisno** iz obeh populacij.

VII. $H_0\ :\ \mu_1-\mu_2=D_0$

Če ne poznamo σ_1 in/ali σ_2, vzorce jemljemo neodvisno, spremenljivki sta vsaka na svoji populacij normalno porazdeljeni, njuni varianci nista enaki, $n_1<30$ ali $n_2<30$, potem

$$\text{T.S.} = \frac{(\overline{y}_1-\overline{y}_2)-D_0}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$ **sledi**

t-porazdelitev z ν prostostnimi stopnjami,

kjer je

$$\nu=\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}$$

Če ν ni naravno število, zaokroži ν navzdol do najbližjega naravnega števila za uporabo t-tabele.

VIII. $H_0\ :\ \mu_d=D_0$

Če vzorce jemljemo neodvisno in če je $n\geq 30$, potem

$$\text{T.S.} = \frac{\overline{d}-D_0}{\frac{s_d}{\sqrt{n}}}$$ **sledi z-porazdelitev.**

IX. $H_0\ :\ \mu_d=D_0$

Če vzorce ne jemljemo neodvisno, če je populacija razlik normalno porazdeljena in če je $n\leq 30$, potem

$$\text{T.S.} = \frac{\overline{d}-D_0}{\frac{s_d}{\sqrt{n}}}$$ **sledi t-porazdelitev**

z $n-1$ prostostnimi stopnjami.

naloga	clovek. urnik	avtomatizirana metoda
1	185,4	180,4
2	146,3	248,5
3	174,4	185,5
4	184,9	216,4
5	240,0	269,3
6	253,8	249,6
7	238,8	282,0
8	263,5	315,9

naloga	clovek. urnik	avtomatizirana metoda	razlika
1	185,4	180,4	5,0
2	146,3	248,5	-102,2
3	174,4	185,5	-11,1
4	184,9	216,4	-31,5
5	240,0	269,3	-29,3
6	253,8	249,6	-4,2
7	238,8	282,0	-43,2
8	263,5	315,9	-52,4

Vstavimo podatke v Minitab (Ex9-40.MTV)

C1: 185,4 146,3 174,4 184,9 240,0 253,8 238,8 263,5

C2: 180,4 248,5 185,5 216,4 269,3 249,6 282,0 315,9

Test za parjenje in interval zaupanja

Parjen T za C1-C2

	N	povpr.	StDev	SE povpr.
C1	8	210,9	43,2	15.3
C2	8	243,4	47,1	16,7
Razlika	8	032,6	35,0	12,4

95% interval zaupanja za razliko povprečja: $(-61,9; -3,3)$

T-test za razliko povpr. = 0 (proti $\neq$ 0):
T-vrednost=-2,63
P-vrednost=0,034.

X. $H_0\ :\ p = p_0$

Če je n dovolj velik, potem

$$\text{T.S.} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} \quad \textbf{sledi } z\textbf{-porazdelitev.}$$

Kot splošno pravilo bomo zahtevali, da velja

$$n\hat{p} \geq 4 \quad \text{in} \quad n\hat{q} \geq 4.$$

Primer (C0)

Postavimo domnevo o vrednosti parametra, npr. π – delež enot z določeno lastnostjo na populaciji. Denimo, da je domneva

$$H\ :\ \pi_H = 0,36$$

Tvorimo slučajne vzorce npr. velikosti $n = 900$ in na vsakem vzorcu določimo vzorčni delež p (delež enot z določeno lastnostjo na vzorcu). Ob predpostavki, da je domneva pravilna, vemo, da se vzorčni deleži porazdeljujejo približno normalno

$$N\left(\pi_H, \sqrt{\frac{\pi_H(1-\pi_H)}{n}}\right)$$

Vzemimo en slučajni vzorec z vzorčnim deležem p. Ta se lahko bolj ali manj razlikuje od π_H. Če se zelo razlikuje, lahko podvomimo o resničnosti domneve π_H. Zato okoli π_H naredimo območje sprejemanja domneve in izven tega območja območje zavračanja domneve.

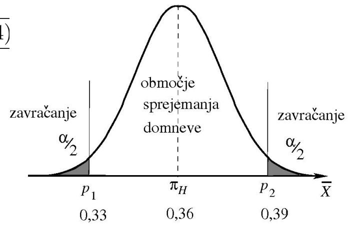
Denimo, da je območje zavračanja določeno s 5% vzorcev,
ki imajo ekstremne vrednosti deležev (2,5% levo in 2,5% desno).

Deleža, ki ločita območje sprejemanja
od območja zavračanja lahko izračunamo takole:

$$p_{1,2} = \pi_H \pm z_{\alpha/2}\sqrt{\frac{\pi_H(1-\pi_H)}{n}},$$

$$p_{1,2} = 0,36 \pm 1,96\sqrt{\frac{0,36\times 0,64)}{900}}$$

$$= 0,36 \pm 0,03$$



Kot smo že omenili, je sprejemanje ali zavračanje domnev po opisanem postopku lahko napačno v dveh smislih:

Napaka 1. vrste (α):

Če vzorčna vrednost deleža pade v območje zavračanja, domnevo π_H zavrnemo. Pri tem pa vemo, da ob resnični domnevi π_H obstajajo vzorci, ki imajo vrednosti v območju zavračanja.
Število α je verjetnost, da vzorčna vrednost pade v območje zavračanja ob prodpostavki, da je domneva resnična.
Zato je α verjetnost, da zavrnemo pravilno domnevo – **napaka 1. vrste**.
Ta napaka je merljiva in jo lahko poljubno manjšamo.

Napaka 2. vrste (β):

Vzorčna vrednost lahko pade v območje sprejemanja,
čeprav je domnevna vrednost parametra napačna.

V primeru, ki ga obravnavamo,
naj bo prava vrednost deleža na populaciji $\pi = 0,40$.

Tedaj je porazdelitev vzorčnih deležev

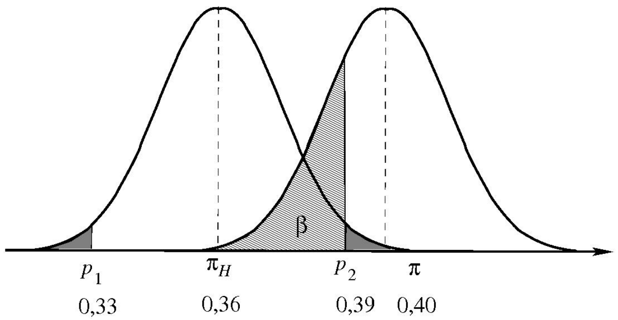
$$N\left(\pi, \sqrt{\frac{\pi(1-\pi)}{n}}\right) = N(0,40; 0,0163)$$

Ker je območje sprejemanja, domneve v intervalu $0,33 < p < 0,39$,
lahko izračunamo verjetnost, da bomo sprejeli napačno domnevo takole:

$$\beta = P(0,33 < p < 0,39) = 0,27$$

Napako 2. vrste lahko izračunamo le,
če imamo znano resnično vrednost parametra π.

Ker ga ponavadi ne poznamo, tudi ne poznamo napake 2. vrste.
Zato ne moremo sprejemati domnev.



Primer (D)

Državni zapisi indicirajo, da je od vseh vozil, ki gredo skozi testiranje izpušnih plinov v preteklem letu, 70% uspešno opravilo testiranje v prvem poskusu.

Naključni vzorec 200ih avtomobilov testiranih v določeni pokrajni v tekočem letu je pokazalo, da jih je 156 šlo čez prvi test.

Ali to nakazuje, da je dejanski delež populacije za to pokrajno v tekočem letu različno od preteklega državnega deleža?

Pri testiranju hipoteze uporabi $\alpha = 0,05$.

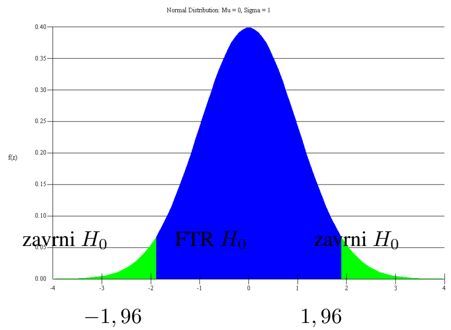
Testiranje hipoteze za delež

- Ničelna hipoteza $H_0\ :\ p = 0,7$
- Alternativna hipoteza $H_a\ :\ p \neq 0,7$
- Test
  - $H_0\ :\ p = p_0$ (delež populacije)
- Predpostavke
  - naključni vzorec
  - izbiranje vzorca iz binomske porazdelitve
- Testna statistika

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}}$$

$$n\hat{p} \geq 4 \quad \text{in} \quad n\hat{q} \geq 4.$$

Določimo zavrnitveni kriterij



Rezultati testiranja

- Naredi naključni vzorec
  - delež vzorca: $156/200 = 0,78$
- Izračunaj vrednost testne statistike

$$Z = (0,78 - 0,7)/0,0324 = 2,4688$$

- Naredi odločitev
  - zavrni hipotezo H_0
- Zaključek
  - Pokrajna ima drugačen kriterij.

P-vrednost

- Sprejemljivost hipoteze H_0 na osnovi vzorca
  - možnost za opazovanje vzorca (ali bolj ekstremno podatkov), če je hipoteza H_0 pravilna
  - P-vrednost $= (2) * P(Z > 2,469) = (2) * (0,0068) = 0,0136$
- Najmanjši α pri katerem zavrnemo hipotezo H_0
  - P-vrednost $< \alpha$, zato zavrni hipotezo H_0

Razlika deležev dveh populaciji

Velik vzorec za testiranje hipoteze o $p_1 - p_2$

Kot splošno pravilo bomo zahtevali, da velja

$$n_1\hat{p}_1 \geq 4, \qquad n_1\hat{q}_1 \geq 4,$$

$$n_2\hat{p}_2 \geq 4 \quad \text{in} \quad n_2\hat{q}_2 \geq 4.$$

XI. Velik vzorec za testiranje hipoteze o $p_1 - p_2$, kadar je $D_0 = 0$.

$$\text{T.S.} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$ **sledi z-porazdelitev.**

kjer je $\hat{p} = \dfrac{y_1 + y_2}{n_1 + n_2}$.

Primer (D3)

Želimo preveriti, ali je predsedniški kandidat različno priljubljen med mestnimi in vaškimi prebivalci. Zato smo slučajni vzorec mestnih prebivalcev povprašali, ali bi glasovali za predsedniškega kandidata.

Od 300 vprašanih (n_1) jih je 90 glasovalo za kandidata (k_1).
Od 200 slučajno izbranih vaških prebivalcev (n_2) pa je za kandidata glasovalo 50 prebivalcev (k_2).

Domnevo, da je kandidat različno priljubljen v teh dveh območjih preverimo pri 10% stopnji značinosti.

$$H_0 \,:\, \pi_1 = \pi_2 \quad \text{oziroma} \quad \pi_1 - \pi_2 = 0,$$

$$H_1 \,:\, \pi_1 \neq \pi_2 \quad \text{oziroma} \quad \pi_1 - \pi_2 \neq 0.$$

Vemo, da se razlika vzorčnih deležev porazdeljuje približno normalno:

$$p_1 - p_2 \,:\, N\left(\pi_1 - \pi_2, \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}\right).$$

Seveda π_1 in π_2 nista znana. Ob predpostavki, da je ničelna domneva pravilna, je matematično upanje razlike vzorčnih deležev hipotetična vrednost razlike deležev, ki je v našem primeru enaka 0. Problem pa je, kako oceniti standardni odklon. Ker velja domneva $\pi_1 = \pi_2 = \pi$, je disperzija razlike vzorčnih deležev

$$\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2} = \frac{\pi(1-\pi)}{n_1} + \frac{\pi(1-\pi)}{n_2} = \pi(1-\pi)\left(\frac{1}{n_1} + \frac{1}{n_2}\right).$$

Populacijski delež π ocenimo z obteženim povprečjem vzorčnih deležev p_1 in p_2

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{k_1 + k_2}{n_1 + n_2}.$$

Vrnimo se na primer. Vzorčna deleža sta:

$$p_1 = \frac{90}{300} = 0,30, \qquad p_2 = \frac{50}{200} = 0,25.$$

Ocena populacijskega deleža je

$$P = \frac{50 + 90}{200 + 300} = 0,28.$$

Kot smo že omenili, se statistika

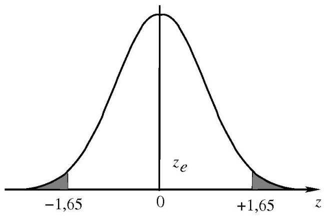
$$z = \frac{p_1 - p_2 - (\pi_1 - \pi_2)_H}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}.$$

porazdeljuje približno standardizirano normalno $N(0,1)$.

Ker gre za dvostranski test, sta kritični vrednosti $\pm z_{\alpha/2} = \pm 1,65$.

Eksperimentalna vrednost statistike pa je

$$z_e = \frac{0,30 - 0,025 - 0}{\sqrt{0,28(1-0,28)\left(\frac{1}{300} + \frac{1}{200}\right)}} = 1,22.$$



Eksperimentalna vrednost ne pade v kritično območje. Zato ničelne domneve ne moremo zavrniti. Priljubljenost predsedniškega kandidata ni statistično značilno različna med mestnimi in vaškimi prebivalci.

XII. Velik vzorec za testiranje hipoteze o $p_1 - p_2$, kadar je $D_0 \neq 0$.

$$\text{T.S.} = \frac{(\hat{p}_1 - \hat{p}_2) - D_0}{\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}}$$ **sledi z-porazdelitev.**

Primer

Neka tovarna cigaret proizvaja dve znamki cigaret. Ugotovljeno je, da ima 56 od 200 kadilcev raje znamko A in da ima 29 od 150 kadilcev raje znamko B.

Testiraj hipotezo pri $0,06$ stopnji zaupanja, da bo prodaja znamke A boljša od prodaje znamke B za 10% proti alternativni hipotezi, da bo razlika manj kot 10%.

Analiza variance

Če opravljamo isti poskus v nespremenjenih pogojih, kljub temu v rezultatu poskusa opažamo spremembe (variacije) ali odstopanja.

Ker vzrokov ne poznamo in jih ne moremo kontrolirati, spremembe pripisujemo *slučajnim vplivom* in jih imenujemo *slučajna odstopanja*.

Če pa enega ali več pogojev v poskusu spreminjamo, seveda dobimo dodatna odstopanja od povprečja. Analiza tega, ali so odstopanja zaradi sprememb različnih faktorjev ali pa zgolj slučajna, in kateri faktorji vplivajo na variacijo, se imenuje *analiza variance*.

Zgleda:

(a) Namesto dveh zdravil proti nespečnosti kot v Studentovem primeru lahko preskušamo učinkovitost več različnih zdravil A, B, C, D,... in s preskušanjem hipoteze $H_0 : \mu_1 = \mu_2 = \cdots = \mu_r$ raziskujemo, ali katero od zdravil sploh vpliva na rezultat. Torej je to posplošitev testa za $H_0 : \mu_1 = \mu_2$

(b) Raziskujemo hektarski donos pšenice. Nanj vplivajo različni faktorji: različne sorte pšenice, različni načini gnojenja, obdelave zemlje itd., nadalje klima, čas sejanja itd .

Analiza variance je nastala prav v zvezi z raziskovanjem v kmetijstvu. Glede na število faktorjev, ki jih spreminjamo, ločimo t.i. *enojno klasifikacijo* ali *enofaktorski eksperiment*, *dvojno klasifikacijo* ali *dvofaktorski eksperiment*, itd.

. . . Analiza variance

V izrazu
$$Q_v^2 = \sum_{i=1}^r \sum_{k=1}^{n_i} (X_{ik} - \overline{X}_i)^2 = \sum_{i=1}^r (n_i - 1) S_i^2$$
je
$$S_i^2 = \frac{1}{n_i - 1} \sum_{k=1}^{n_i} (X_{ik} - \overline{X}_i)^2$$
nepristranska cenilka za disperzijo v i-ti skupini; neodvisna od S_j^2, za $i \neq j$. Zato ima
$$\frac{Q_v^2}{\sigma^2} = \sum_{i=1}^r (n_i - 1) \frac{S_i^2}{\sigma^2}$$
porazdelitev $\chi^2(n-r)$, saj je ravno $\sum_{i=1}^r (n_i - 1) = n - r$ prostostnih stopenj.

Ker je $\mathsf{E}\frac{Q_v^2}{\sigma^2} = n - r$, je tudi $S_v^2 = \dfrac{1}{n-r} Q_v^2$ nepristranska cenilka za σ^2.

. . . Analiza variance

Izračunajmo še Q_m^2 pri predpostavki o veljavnosti osnovne domneve H_0. Dobimo
$$Q_m^2 = \sum_{i=1}^r n_i (\overline{X}_i - \mu)^2 - n(\overline{X} - \mu)^2.$$
Torej je
$$\frac{Q_m^2}{\sigma^2} = \sum_{i=1}^r n_i \Big(\frac{\overline{X}_i - \mu}{\sigma/\sqrt{n_i}} \Big)^2 - n \Big(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \Big)^2$$
od tu spregvidimo, da je statistika $\frac{Q_m^2}{\sigma^2}$ porazdeljena po $\chi^2(r-1)$.

Poleg tega je $S_m^2 = \dfrac{Q_m^2}{r-1}$ nepristranska cenilka za σ^2, neodvisna od S_v^2.

. . . Analiza variance

Ker sta obe cenilki za varianco σ^2, pri domnevi H_0, njuno razmerje $F = \frac{S_m^2}{S_v^2}$ ne more biti zelo veliko. Iz
$$F = \frac{S_m^2}{S_v^2} = \frac{Q_m^2/(r-1)}{Q_v^2/(n-r)} = \frac{\frac{Q_m^2}{\sigma^2}/(r-1)}{\frac{Q_v^2}{\sigma^2}/(n-r)}$$
vidimo da gre za Fisherjevo (Snedecorjevo) porazdelitev $F(r-1, n-r)$.

Podatke zapišemo v *tabelo analize variance*

VV	VK	PS	PK	F
faktor	Q_m^2	$r-1$	S_m^2	F
slučaj	Q_v^2	$n-r$	S_v^2	
	Q^2	$n-1$		

Analiza variance v R-ju

Zgled: Petnajst enako velikih njiv je bilo posejanih z isto vrsto pšenice, vendar gnojeno na tri različne načine – z vsakim po pet njiv.

```
> ena <- c(47,47,40,32,40)
> dva <- c(76,68,71,46,54)
> tri <- c(49,40,34,36,44)
> d <- stack(list(e=ena,d=dva,t=tri))
> names(d)
[1] "values" "ind"
> oneway.test(values ~ ind, data=d, var.equal=TRUE)

        One-way analysis of means

data:  values and ind
F = 10.5092, num df = 2, denom df = 12, p-value = 0.002304
> av <- aov(values ~ ind, data=d)
> summary(av)
            Df  Sum Sq Mean Sq F value   Pr(>F)
ind          2 1628.93  814.47  10.509 0.002304 **
Residuals   12  930.00   77.50
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Domnevo H_0 zavrnemo.

XIII. Testiranje hipoteze o varianci populacije $H_0 : \sigma^2 = \sigma_0^2$

$$\text{T.S.} = \frac{(n-1)s^2}{\sigma_0^2} \ \text{sledi}\ \chi^2\text{−porazd.}$$

Če je

- $H_a : \sigma^2 > \sigma_0^2$, potem je **odločitveno pravilo**: zavrni ničelno hipotezo, če je test statistike večji ali enak $\chi^2_{(\alpha,n-1)}$.
- $H_a : \sigma^2 < \sigma_0^2$ potem je **odločitveno pravilo**: zavrni ničelno hipotezo, če je test statistike manjši ali enak $\chi^2_{(\alpha,n-1)}$.
- $H_a : \sigma^2 \neq \sigma_0^2$, potem je **odločitveno pravilo**: zavrni ničelno hipotezo, če je test statistike manjši ali enak $\chi^2_{(\alpha,n-1)}$ ali če je test statistike večji ali enak $\chi^2_{(\alpha,n-1)}$.

Primer (E)

Količina pijače, ki jo naprava za mrzle napitke zavrže je normalno porazdeljena s povprečjem 12 unčev in standardnim odklonom 0,1 unče.

Vsakič, ko servisirajo napravo, si izberejo 10 vzorcev in izmerijo zavrženo tekočino.

Če je razpršenost zavržene količine prevelika, potem mora naprava na servis.

Ali naj jo odpeljejo na servis?

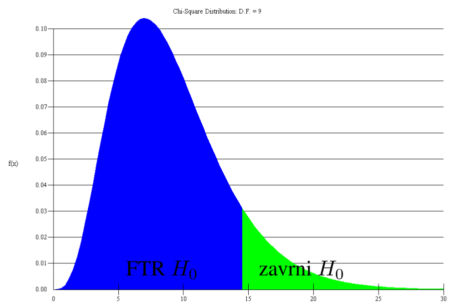
Uporabi $\alpha = 0,1$.

Testiranje hipoteze za varianco

- Ničelna hipoteza $H_0 : \sigma^2 = 0,01$,
- Alternativna hipoteza $H_a : \sigma^2 > 0,01$,
- Predpostavke
  - naključni vzorec
  - vzorčenje iz normalne porazdelitve.
- Testna statistika

$$\chi^2_{\nu=n-1} = \frac{S^2(n-1)}{\sigma_0^2}.$$

Določimo zavrnitveni kriterij



14,684

Rezultati testiranja

- naredi naključni vzorec
  - varianca vzorca: $0,02041$
- izračunaj vrednost testne statistike

$$\chi^2 = (0,02041)(9)/(0,01) = 18,369$$

- naredi odločitev
  - zavrni H_0
- zaključek
  - popravi napravo

P-vrednost

- Sprejemljivost hipoteze H_0 na osnovi vzorca
  - možnost za opazovanje vzorca (ali bolj ekstremno podatkov), če je hipoteza H_0 pravilna
  - P-vrednost $= P(\chi^2 > 18,369) = 0,0311$
- Najmanjši α pri katerem zavrnemo hipotezo H_0
  - P-vrednost $< \alpha$, zato zavrni hipotezo H_0.

XIV. Testiranje hipoteze o kvocientu varianc neodvisnih vzorcev $H_0 : \sigma_1^2/\sigma_2^2 = 1$

Če velja

$$H_a : \sigma_1^2/\sigma_2^2 > 1,$$

potem je **test statistike** enak s_1^2/s_2^2,

odločitveno pravilo pa je:

zavrni ničelno hipotezo, če velja

$$\text{T.S.} \geq F_{\alpha, n_1-1, n_2-1}.$$

Če velja $H_a : \sigma_1^2/\sigma_2^2 < 1$, potem je **test statistike** enak

$$\frac{\text{varianca večjega vzorca}}{\text{varianca manjšega vzorca}}.$$

odločitveno pravilo pa je:

zavrni ničelno hipotezo, če velja $s_1^2 > s_2^2$ in

$$\text{T.S.} \geq F_{\alpha, n_1-1, n_2-1}$$

oziroma zavrni ničelno hipotezo, če velja $s_1^2 < s_2^2$ in

$$\text{T.S.} \geq F_{\alpha, n_2-1, n_1-1}.$$