STATISZTIKA 2.

KÉPLETGYŰJTEMÉNY

matekking.hu

időssorok
statisztikai becslések
hipotézisvizsgálat
regressziószámítás

6. IDŐSOROK

6.1. Állapotidősor és tartamidősor

	ÁLLAPOTIDŐSOR	TARTAMIDŐSOR
Változás mértéke	$\bar{d}=\frac{\sum d_t}{n-1}=\frac{y_n-y_1}{n-1}$	$\bar{d}=\frac{\sum d_t}{n-1}=\frac{y_n-y_1}{n-1}$
Változás üteme	$\bar{l}=\sqrt[n-1]{\prod_{t=2}^{n} l_t}=\sqrt[n-1]{\frac{y_n}{y_1}}$	$\bar{l}=\sqrt[n-1]{\prod_{t=2}^{n} l_t}=\sqrt[n-1]{\frac{y_n}{y_1}}$
Átlag	$\bar{y}_k=\frac{\frac{y_1}{2}+y_2+\ldots+\frac{y_n}{2}}{n-1}$	$\bar{y}=\frac{y_1+y_2+\ldots+y_n}{n}$

6.2. Mozgóátlagok

Ha a tagok száma páratlan:

$$\hat{y}_t=\frac{y_{t-k}+\ldots+y_{t-1}+y_t+y_{t+1}+\ldots+y_{t+k}}{2k+1}$$

Ha pedig a tagok száma páros

$$\hat{y}_t=\frac{\frac{y_{t-k}}{2}+\ldots+y_{t-1}+y_t+y_{t+1}+\ldots+\frac{y_{t+k}}{2}}{2k}$$

6.3. Lineáris és exponenciális trend

Lineáris trend

$$\hat{y}=\beta_0+\beta_1\cdot t$$

Lineáris trend normálegyenletei

$$\sum_{t=1}^{n} y_t=n\cdot\beta_0+\beta_1\sum_{t=1}^{n} t \qquad \sum_{t=1}^{n} t\cdot y_t=\beta_0\cdot\sum_{t=1}^{n} t+\beta_1\sum_{t=1}^{n} t^2$$

Exponenciális trend

$$\hat{y}=\beta_0\cdot\beta_1^{\,t}$$

$$\ln\hat{y}=\ln\beta_0+t\cdot\ln\beta_1$$

info@mateking.hu
tel:06705411417
mateking.hu

6.4. Szezonális eltérés lineáris trend esetén

$$s_j = \frac{\sum_{i=1}^{n/p}\left(y_{ij} - \hat{y}_{ij}\right)}{n/p}$$

ÉVEK=i	SZEZONOK=j (szezonfajták száma p)			
	j=1	j=2	j=3	…
i=1	$y_{1\,1}$	$y_{1\,2}$	$y_{1\,3}$	$y_{1\,4}$
i=2	$y_{2\,1}$	$y_{2\,2}$	$y_{2\,3}$	$y_{2\,4}$
i=3	$y_{3\,1}$	$y_{3\,2}$	$y_{3\,3}$	$y_{3\,4}$
…	$y_{4\,1}$	$y_{4\,2}$	$y_{4\,3}$	$y_{4\,4}$

6.5. Korrigált szezonális eltérés lineáris trend esetén

$$\breve{s}_j = s_j - \bar{s}$$

6.6. Szezonindex exponenciális trend esetén

$$s^*_j = \frac{\sum_{i=1}^{n/p}\left(\frac{y_{ij}}{\hat{y}_{ij}}\right)}{n/p}$$

mateking.hu

6.7. Korrigált szezonindex exponenciális trend esetén

$$\breve{s}^*_j = \frac{s^*_j}{\bar{s}}$$

7. STATISZTIKAI BECSLÉSEK

7.1. Becslések

Torzítatlanság: $E(\hat{\theta}) = \theta$

$$MSE(\hat{\theta}) = \mathrm{var}(\hat{\theta}) + \left(E(\hat{\theta}) - \hat{\theta}\right)^2$$

7.2. Sokasági átlag, arány és variancia intervallumbecslése FAE-minták esetén

ÁTLAG INTERVALLUMBECSLÉSE, HA A SOKASÁGI SZÓRÁS ISMERT (FAE MINTA)

$$\bar{x} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$1-\alpha =$ konfidencia szint
$\bar{x}$ = a minta átlaga
n = a minta elemszáma
σ = a teljes sokaság szórása
$Z_{1-\frac{\alpha}{2}}$ = a standard normális eloszlás $1-\frac{\alpha}{2}$ valószínűséghez tartozó Z értéke, lásd táblázat

ARÁNY INTERVALLUMBECSLÉSE (FAE MINTA)

$$\bar{p} \pm Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

$1-\alpha =$ konfidencia szint
$\bar{p} =$ a minta alapján kapott valószínűség
$n =$ a minta elemszáma.
$Z_{1-\frac{\alpha}{2}} =$ a standard normális eloszlás $1-\frac{\alpha}{2}$ valószínűséghez tartozó Z-értéke, lásd táblázat.

ÁTLAG INTERVALLUMBECSLÉSE, HA A SOKASÁGI SZÓRÁS NEM ISMERT (FAE MINTA)

$$\bar{x} \pm t_{1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$$

$1-\alpha =$ konfidencia szint
$\bar{x}$ = a minta átlaga
n = a minta elemszáma
s = a minta szórása, a sokasági szórás nem
$t_{1-\frac{\alpha}{2}}$ = a t-eloszlás $1-\frac{\alpha}{2}$-höz tartozó értéke.

VARIANCIA INTERVALLUMBECSLÉSE (FAE MINTA)

$$\frac{(n-1)\cdot s^2}{\chi^2_{1-\alpha/2}(v)} < \sigma^2 < \frac{(n-1)\cdot s^2}{\chi^2_{\alpha/2}(v)}$$

$1-\alpha =$ konfidencia szint
n = a minta elemszáma
s = a minta szórása, a sokasági szórás nem
$\chi^2(v)$ = a khí-négyzet eloszlás megfelelő értéke

info@mateking.hu
tel:06705411417

7.3. Sokasági átlag, arány intervallumbecslése EV-mintából

ÁTLAG INTERVALLUMBECSLÉSE, HA A SOKASÁGI SZÓRÁS ISMERT (EV-MINTA)

$$\bar{x} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \cdot \sqrt{1-\frac{n}{N}}$$

$1-\alpha =$ konfidencia szint
$\bar{x}$ = a minta átlaga
n = a minta elemszáma
N = a teljes sokaság elemszáma
σ = a teljes sokaság szórása
$Z_{1-\frac{\alpha}{2}}$ = a standard normális eloszlás $1-\frac{\alpha}{2}$ valószí-

ARÁNY INTERVALLUMBECSLÉSE (EV-MINTA)

$$\bar{p} \pm Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \cdot \sqrt{1-\frac{n}{N}}$$

$1-\alpha =$ konfidencia szint
$\bar{p} =$ a minta alapján kapott valószínűség
$n =$ a minta elemszáma
N = a teljes sokaság elemszáma

$Z_{1-\frac{\alpha}{2}}$ = a standard normális eloszlás $1-\frac{\alpha}{2}$ valószí-

7.4. Kétmintás becslések

KÉT ÁTLAG KÜLÖNBSÉGÉNEK BECSLÉSE

$$d \pm t_{1-\frac{\alpha}{2}} \cdot s_d \quad \text{ahol} \quad d = \bar{X} - \bar{Y}$$

$$s_d = s_c \cdot \sqrt{\frac{1}{n_Y} + \frac{1}{n_X}} \quad \text{itt} \quad s_c^2 = \frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2}$$

$1-\alpha =$ konfidencia szint
$\bar{x}$ = az egyik minta átlaga
$\bar{Y}$ = az másik minta átlaga
n_X = az egyik minta elemszáma
n_Y = a másik minta elemszáma
A szabadságfok $v = n_X + n_Y - 2$

HÁNYADOSBECSLÉS

$$H = \frac{\bar{Y}}{\bar{X}} \quad \text{ahol} \quad h_1 = \frac{\bar{y}}{\bar{x}} \quad \text{és} \quad E(h_1) \approx H + \frac{H}{n}\left[V_x^2 - rV_xV_y\right]$$

$$Var(h_1) \approx \frac{H^2}{n}\left[V_x^2 + V_y^2 - 2rV_xV_y\right] \qquad \hat{\bar{Y}}_H = \bar{X} \cdot \frac{\bar{y}}{\bar{x}} = \bar{X}h_1$$

7.5. Rétegzett minták

ÁTLAG INTERVALLUMBECSLÉSE RÉTEGZETT MINTÁBÓL

$$\bar{x}_R \pm z_{1-\frac{\alpha}{2}} \cdot s_{\hat{\bar{X}}_R} \quad \text{ahol} \quad s_{\hat{\bar{X}}_R} = \sum_{j=1}^{M} W_j^2 \frac{s_j^2}{n_j}\left(1-\frac{n_j}{N_j}\right)$$

$1-\alpha=$konfidencia szint
$\bar{x}$ = a minta átlaga
n = a minta elemszáma
n_j = a minta j-edik rétegének elemszáma
N = a teljes sokaság elemszáma
N_j = a teljes sokaság j-edik rétegének elemszáma
W_j = a teljes sokaság j-edik rétegének a teljes sokasághoz viszonyított aránya
s_j = a minta j-edik rétegének szórása

$Z_{1-\frac{\alpha}{2}}$ =a standard normális eloszlás $1-\frac{\alpha}{2}$ valószínűséghez tartozó Z értéke.

mateking.hu

7.6. Független részminták módszere

$$\hat{\theta}_{FRM} = \frac{\sum_{i=1}^{k} \hat{\theta}_i(m)}{k}$$

$$\operatorname{var}(\hat{\theta}_{FRM}) = \frac{\sum_{i=1}^{k} \left(\hat{\theta}_i(m) - \hat{\theta}_{FRM}\right)^2}{k(k-1)}$$

$$s_{\hat{\theta}_{FRM}} = \sqrt{\operatorname{var}(\hat{\theta}_{FRM})}$$

$$t = \frac{\hat{\theta}_{FRM} - \theta}{s_{\hat{\theta}_{FRM}}}$$

8. HIPOTÉZISVIZSGÁLAT

Minta	Paraméteres Próbák			Nemparaméteres próbák		
	Átlag	**Arány**	**Szórás**	**Eloszlás**	**Függetlenség**	
Egy mintás	Aszimptotikus Z-próba	Z-próba		Illeszkedés-vizsgálat: χ^2-próba	Függetlenség-vizsgálat: χ^2-próba	NAGY MINTA, BÁRMILYEN ELOSZLÁSÚ SOKASÁGRA
	Z-próba t-próba		χ^2-próba			CSAK NORMÁLIS ELOSZLÁSÚ SOKASÁGRA
Két mintás	Aszimptotikus Z-próba	Z-próba		Homogenitás-vizsgálat: χ^2-próba		NAGY MINTA, BÁRMILYEN ELOSZLÁSÚ SOKASÁGRA
	Z-próba t-próba		F-próba			CSAK NORMÁLIS ELOSZLÁSÚ SOKASÁGRA
Több mintás	Variancia-analízis		Bartlett			

A statisztikai próbákat két nagy típusba sorolhatjuk. Vannak az úgynevezett paraméteres próbák, amik egy sokaság – esetleg több sokaság – valamilyen paraméterével kapcsolatos hipotézissel foglalkoznak. Ilyen paraméter tipikusan az átlag a szórás és az arány, de természetesen bármilyen más paramétert is vizsgálhatunk. A másik nagy csoport a nemparaméteres próbák, amik a sokaság eloszlására, vagy a sokaságon belüli ismérvek eloszlásának egyezőségére, esetleg azok függetlenségére irányuló hipotézisekkel kapcsolatos próbák.
A paraméteres és a nemparaméteres próbákon belül megkülönböztetünk egy mintás, két mintás és több mintás próbákat.

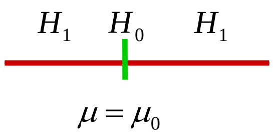
8.1. Z-próba

Kritikus értékek α szignifikanciaszint esetén

KÉTOLDALI KRITIKUS TARTOMÁNY

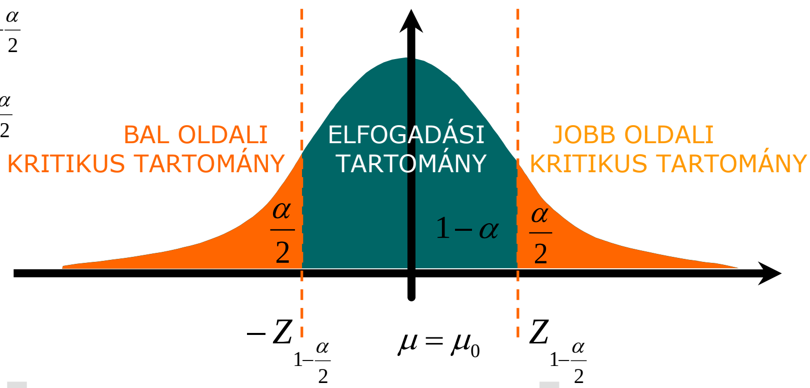
$H_0 : \mu = \mu_0$

$H_1 : \mu \neq \mu_0$



BAL OLDALI KRITIKUS ÉRTÉK: $-Z_{1-\frac{\alpha}{2}}$

JOBB OLDALI KRITIKUS ÉRTÉK: $Z_{1-\frac{\alpha}{2}}$

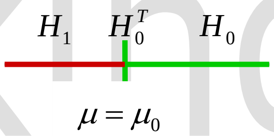


mateking.hu

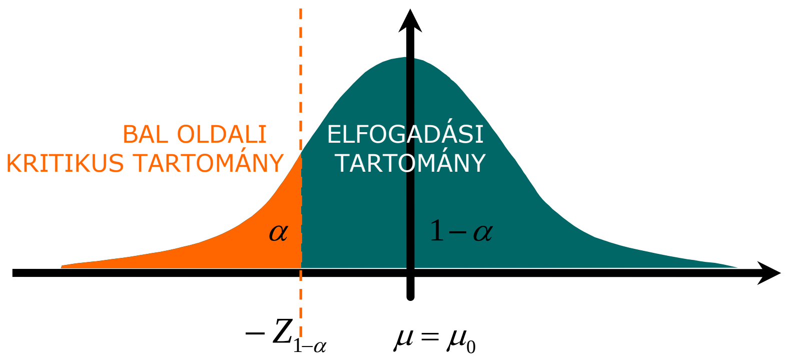
BAL OLDALI KRITIKUS TARTOMÁNY

$H_0 : \mu \geq \mu_0 \qquad H_0^T : \mu = \mu_0$

$H_1 : \mu < \mu_0$



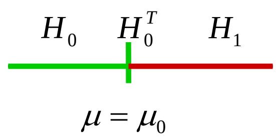
BAL OLDALI KRITIKUS ÉRTÉK: $-Z_{1-\alpha}$



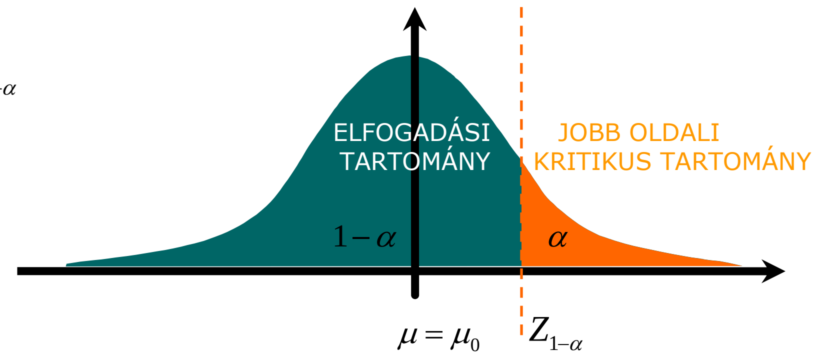
JOBB OLDALI KRITIKUS TARTOMÁNY

$H_0 : \mu \leq \mu_0 \qquad H_0^T : \mu = \mu_0$

$H_1 : \mu > \mu_0$



JOBB OLDALI KRITIKUS ÉRTÉK: $Z_{1-\alpha}$



© www.mateking.hu
info@mateking.hu
tel:06705411417

8.1.1. Sokasági átlagra vonatkozó hipotézis, Z-próba

Z-próba: A sokaság normális eloszlású, szórása σ , H_0 a sokaság átlagára vonatkozik, a minta elemszáma n.

$$Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

8.1.2. Sokasági átlagra vonatkozó hipotézis, aszimptotikus Z-próba

Aszimptotikus Z-próba: A sokaság tetszőleges eloszlású, szórása nem ismert, H_0 a sokaság átlagára vonatkozik, a minta n elemű, elemszáma nagy.

$$Z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

8.1.3. Sokasági arányra vonatkozó hipotézis, Z-próba

Z-próba: A sokaság tetszőleges eloszlású, H_0 egy sokasági arányra vonatkozik, a minta n elemű, elemszáma nagy.

$$Z = \frac{P - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}}$$

8.1.4. Sokasági átlagra vonatkozó hipotézis, t-próba

t-próba: A sokaság normális eloszlású, szórása nem ismert, H_0 a sokaság átlagára vonatkozik, a minta elemszáma n.

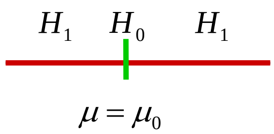
$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Szabadságfok: v=n-1. Kritikus értékek α szignifikanciaszint esetén

KÉTOLDALI KRITIKUS TARTOMÁNY

$H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$

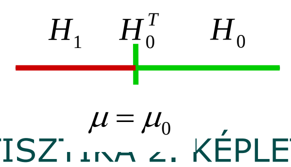


BAL OLDALI KRITIKUS ÉRTÉK: $-t_{1-\frac{\alpha}{2}}$ JOBB OLDALI KRITIKUS ÉRTÉK: $t_{1-\frac{\alpha}{2}}$

BAL OLDALI KRITIKUS TARTOMÁNY

$H_0: \mu \geq \mu_0$ $H_0^T: \ \mu = \mu_0$

$H_1: \mu < \mu_0$



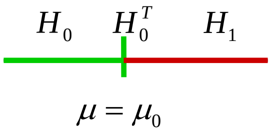
BAL OLDALI KRITIKUS ÉRTÉK: $-t_{1-\alpha}$

info@mateking.hu
tel:06705411417

JOBB OLDALI KRITIKUS TARTOMÁNY

$H_0: \mu \le \mu_0 \qquad H_0^T: \mu = \mu_0$

$H_1: \mu > \mu_0$



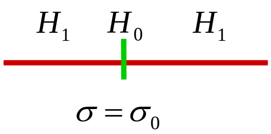
JOBB OLDALI KRITIKUS ÉRTÉK: $t_{1-\alpha}$

8.2. χ^2-próba

KÉTOLDALI KRITIKUS TARTOMÁNY

$H_0: \sigma = \sigma_0$

$H_1: \sigma \ne \sigma_0$



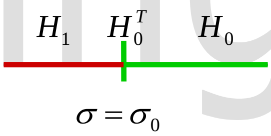
BAL OLDALI KRITIKUS ÉRTÉK: $\chi^2_{\frac{\alpha}{2}}$

JOBB OLDALI KRITIKUS ÉRTÉK: $\chi^2_{1-\frac{\alpha}{2}}$

BAL OLDALI KRITIKUS TARTOMÁNY

$H_0: \sigma \ge \sigma_0 \qquad H_0^T: \sigma = \sigma_0$

$H_1: \sigma < \sigma_0$

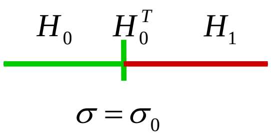


BAL OLDALI KRITIKUS ÉRTÉK: χ^2_{α}

JOBB OLDALI KRITIKUS TARTOMÁNY

$H_0: \sigma \le \sigma_0 \qquad H_0^T: \sigma = \sigma_0$

$H_1: \sigma > \sigma_0$



JOBB OLDALI KRITIKUS ÉRTÉK: $\chi^2_{1-\alpha}$

8.2.1. Sokasági varianciára vonatkozó hipotézis, χ^2-próba

χ^2-próba: A sokaság normális eloszlású, H_0 a sokasági szórásra vonatkozik, a minta n elemű.

$$\chi^2 = \frac{(n-1)\cdot s^2}{\sigma_0^2}$$

Szabadságfok: v=n-1

mateking.hu

8.3. Nemparaméteres próbák

8.3.1. Illeszkedésvizsgálat, χ^2-próba

A sokaság eloszlására irányuló vizsgálat, H_0: mindegyik osztályköz valószínűsége egy adott eloszlásnak megfelelő érték, vagyis minden i-re az i-edik osztályköz valószínűsége a P_i érték. Az ellenhipotézis H_1: van olyan osztályköz, ami nem az adott eloszlásnak megfelelő P_i érték.

A próbát $\chi^2_{1-\alpha}(v)$ jobb oldali kritikus értékkel végezzük el, a nullhipotézist az ennél kisebb, az ellenhipotézist az ennél nagyobb értékek igazolják. A minta elemszáma n.

$$\chi^2(v) = \sum_{i=1}^{k} \frac{(f_i - np_i)^2}{nP_i}$$

ahol a v szabadságfok $v = k - b - 1$.
Itt $k =$ az osztályközök száma és $b =$ az adott eloszlás azon paramétereinek száma, amit a mintából becsléssel határozunk meg

8.3.2. Függetlenségvizsgálat, χ^2-próba

A sokaságon belül két ismérv függetlenségére irányuló vizsgálat. H_0: a két ismérv független, az ellenhipotézis pedig, H_1: a két ismérv közti kapcsolat sztochasztikus vagy függvényszerű.

A próbát $\chi^2_{1-\alpha}(v)$ jobb oldali kritikus értékkel végezzük el, a nullhipotézist az ennél kisebb, az ellenhipotézist az ennél nagyobb értékek igazolják. A minta elemszáma n, a minta alapján készített kontingencia tábla sorainak száma r, oszlopainak száma c.

$$\chi^2(v) = \sum \frac{\left(n_{ij} - n_{ij}^*\right)^2}{n_{ij}^*}$$

Ahol a v szabadságfok $v = (r-1)(c-1)$.

8.3.3. Homogenitásvizsgálat, χ^2-próba

Két sokaságban valamely változó eloszlásának egyezőségére irányuló vizsgálat. H_0: a két sokaságban az eloszlás egyező, az ellenhipotézis pedig, H_1: a két eloszlás nem egyező.

A próbát $\chi^2_{1-\alpha}(v)$ jobb oldali kritikus értékkel végezzük el, a nullhipotézist az ennél kisebb, az ellenhipotézist az ennél nagyobb értékek igazolják. Mintát ezúttal mindkét sokaságból veszünk, az X sokaságból vett minta elemszáma n_X az Y sokaságból vett mintáé n_Y mindkét mintában az osztályközök száma k.

$$\chi^2(v) = n_X n_Y \cdot \sum_{i=1}^{k} \frac{1}{n_{Xi} + n_{Yi}} \cdot \left(\frac{n_{Xi}}{n_X} - \frac{n_{Yi}}{n_Y}\right)^2$$

Ahol a v szabadságfok $v = k - 1$.

info@mateking.hu
tel:06705411417

8.4. Kétmintás próbák

8.4.1. Két sokaság átlagának eltérésére vonatkozó hipotézis, Z-próba

Kétmintás Z-próba: Mindkét sokaság normális eloszlású, szórásaik ismertek, σ_X és σ_Y.

$$Z = \frac{(\bar{y}-\bar{x})-\delta_0}{\sqrt{\frac{\sigma_Y^2}{n_Y}+\frac{\sigma_X^2}{n_X}}}$$

A nullhipotézis $H_0: \mu_X - \mu_Y = \delta_0$, ahol δ_0 tetszőleges, de előre megadott érték.

A minták elemszáma n_X és n_Y.

8.4.2. Két sokaság átlagának eltérésére vonatkozó hipotézis, t-próba

Kétmintás t-próba: A két sokaság normális eloszlású és szórásaik egyformák.

$$t(v) = \frac{(\bar{y}-\bar{x})-\delta_0}{s\cdot\sqrt{\frac{1}{n_Y}+\frac{1}{n_X}}} \quad \text{itt} \quad s^2 = \frac{(n_X-1)s_X^2+(n_Y-1)s_Y^2}{n_X+n_Y-2}$$

A nullhipotézis $H_0: \mu_X - \mu_Y = \delta_0$, ahol δ_0 tetszőleges, de előre megadott érték.

A minták elemszáma n_X és n_Y, szórása s_X és s_Y, a szabadságfok $v = n_Y + n_X - 2$

mateking.hu

8.4.3. Két sokaság átlagának eltérésére vonatkozó hipotézis, aszimptotikus Z-próba

Kétmintás aszimptotikus Z-próba: A két sokaság eloszlása és szórása nem ismert, mindkettő szórása véges, és mindkét minta elemszáma elég nagy.

$$Z = \frac{(\bar{y}-\bar{x})-\delta_0}{\sqrt{\frac{s_Y^2}{n_Y}+\frac{s_X^2}{n_X}}}$$

A nullhipotézis $H_0: \mu_X - \mu_Y = \delta_0$, ahol δ_0 tetszőleges, de előre megadott érték.

A minták elemszáma n_X és n_Y, szórása s_X és s_Y.

8.4.4. Két sokaság arányának eltérésére vonatkozó hipotézis, Z-próba

Kétmintás Z-próba: Két sokaság sokasági arányának összehasonlítására irányuló próba.

$$Z = \frac{(P_Y - P_X)-\varepsilon_0}{\sqrt{\frac{P_Y(1-P_Y)}{n_Y}+\frac{P_X(1-P_X)}{n_X}}} \qquad \varepsilon_0 = 0 \text{ speciális esetben } Z_0 = \frac{P_Y - P_X}{\sqrt{\bar{p}\cdot\bar{q}\left(\frac{1}{n_Y}+\frac{1}{n_X}\right)}}$$

ahol n_X és n_Y a minták elemszáma.

A nullhipotézis $H_0: P_X - P_Y = \varepsilon_0$, ahol ε_0 tetszőleges, de előre megadott érték.

Abban az esetben, ha $\varepsilon_0 = 0$ a Z_0 próbafüggvényt célszerű alkalmazni, itt

$$\bar{p} = \frac{n_Y P_Y + n_X P_X}{n_Y + n_X} \quad \text{és} \quad \bar{q} = 1-\bar{p}$$

8.4.5. Két sokaság szórásának eltérésére vonatkozó hipotézis, F-próba

F-próba: Két sokaság szórásának összehasonlítására irányuló próba, ha mindkét sokaság normális eloszlású. A nullhipotézis $H_0: \sigma_1^2 = \sigma_2^2$

$$F = \frac{s_1^2}{s_2^2}$$

az F-eloszlás két szabadságfoka $v_1 = n_1 - 1$ és $v_2 = n_2 - 1$, ahol n_1 és n_2 a két minta elemszáma. Célszerű 1-es sokaságnak mindig a nagyobb szórással rendelkezőt nevezni.

A kritikus értékek az $F_{1-p}(v_1; v_2) = \dfrac{1}{F_p(v_2; v_1)}$ összefüggés alapján:

BAL OLDALI KRITIKUS TARTOMÁNY ESETÉN: $\dfrac{1}{F_{1-\alpha}(v_2; v_1)}$

KÉTOLDALI KRITIKUS TARTOMÁNY ESETÉN: $\dfrac{1}{F_{1-\frac{\alpha}{2}}(v_2; v_1)}$ ÉS $F_{1-\frac{\alpha}{2}}(v_1; v_2)$

JOBB OLDALI KRITIKUS TARTOMÁNY ESETÉN: $F_{1-\alpha}(v_1; v_2)$

mateking.hu

8.5. Többmintás próbák

8.5.1. Varianciaanalízis

Több sokaság várható értékének összehasonlítására vonatkozó próba, ha mindegyik sokaság normális eloszlású és azonos szórású.

A H_0 nullhipotézis: $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_M = \mu$, vagyis az, hogy a várható értékek az összes sokaságra (M db) megegyeznek, míg az ellenhipotézis az, hogy van olyan μ_j amire $\mu_j \neq \mu$.

A részsokaságokból vett minták, a részsokaságok száma M.

minta	elemszám	átlag	szórás
1-es részsokaság	n_1	$\bar{x}_1$	s_1
2-es részsokaság	n_2	$\bar{x}_2$	s_2
j-edik részsokaság	n_j	$\bar{x}_j$	s_j
összesen	n	$\bar{x}$	s

$$SSK = \sum_{j=1}^{M} n_j (\bar{x} - \bar{x}_j)^2$$

$$SSB = \sum_{j=1}^{M} (n_j - 1) s_j^2$$

A próbafüggvény

$$F(v_1; v_2) = \frac{SSK/(M-1)}{SSB/(n-M)}$$

A két szabadságfok $v_1 = M-1$és $v_2 = n-M$, a próba jobb oldali kritikus értékkel hajtandó végre: $F_{1-\alpha}(v_1;v_2)$

VARIANCIAANALÍZIS-TÁBLÁZAT

SZÓRÓDÁS OKA	ELTÉRÉS-NÉGYZETÖSSZEG	SZABADSÁG-FOK	ÁTLAGOS NÉGYZETÖSSZEG	F	p-ÉRTÉK
Részsokaságra bontás miatt	SSK	$M-1$	$s_k^2 = \dfrac{SSK}{M-1}$	$F = \dfrac{s_k^2}{s_b^2}$	p
Részsokaságon belüli hiba	SSB	$n-M$	$s_b^2 = \dfrac{SSB}{n-M}$		
össz.	SST	$n-1$			

8.5.2. Bartlett-próba

Bartlett-próba: Több sokaság szórásának összehasonlítására vonatkozó próba, ha mindegyik sokaság normális eloszlású.

A H_0 nullhipotézis: $\sigma_1 = \sigma_2 = \sigma_3 = \ldots = \sigma_M = \sigma$, vagyis az, hogy az összes sokaság (M db) szórása megegyezik, míg az ellenhipotézis az, hogy van olyan σ_j amire $\sigma_j \neq \sigma$.

mateking.hu

A részsokaságokból vett minták, a részsokaságok száma M.

minta	elemszám	átlag	szórás
1-es részsokaság	n_1	$\bar{x}_1$	s_1
2-es részsokaság	n_2	$\bar{x}_2$	s_2
j-edik részsokaság	n_j	$\bar{x}_j$	s_j
összesen	n	$\bar{x}$	s

$$SSK = \sum_{j=1}^{M} n_j (\bar{x} - \bar{x}_j)^2$$

$$SSB = \sum_{j=1}^{M} n_j s_j^2$$

A próbafüggvény

$$B^2 = \frac{1}{c}\left(v \cdot \ln s_b^2 - \sum_{j=1}^{M} v_j \ln s_j^2 \right)$$

A próbafüggvény M-1 szabadságfokú χ^2 eloszlást követ.

v_j a j-edik részsokaság szabadságfoka, tehát $v_j = n_j - 1$ és $v = \sum_{j=1}^{M} v_j = n - M$

$c = 1 + \dfrac{1}{3(M-1)}\left(\sum_{j=1}^{M} \dfrac{1}{v_j} - \dfrac{1}{v} \right)$ $s_b = \dfrac{SSB}{n-M}$ és s_j pedig a részsokaságok szórásai.

A próba jobb oldali kritikus értékkel hajtandó végre: $\chi^2_{1-\alpha}(M-1)$

A próbák áttekintése és a kritikus értékek

Minta	Paraméteres próbák			Nemparaméteres próbák		
	Átlag	**Arány**	**Szórás**	**Eloszlás**	**Függetlenség**	
Egy mintás	Aszimptotikus Z-próba	Z-próba		Illeszkedés-vizsgálat: χ^2-próba	Függetlenség-vizsgálat: χ^2-próba	NAGY MINTA, BÁRMILYEN ELOSZLÁSÚ SOKASÁGRA
	Z-próba t-próba		χ^2-próba			CSAK NORMÁLIS ELOSZLÁSÚ SOKASÁGRA
	KÉTOLDALI $C_a=-Z_{1-\frac{\alpha}{2}}$ $C_f=-Z_{1-\frac{\alpha}{2}}$ BAL OLDALI $C_a=-Z_{1-\alpha}$ JOBB OLDALI $C_f=Z_{1-\alpha}$		KÉTOLDALI $C_a=\chi^2_{\frac{\alpha}{2}}$ $C_f=\chi^2_{1-\frac{\alpha}{2}}$ BAL OLDALI $C_a=\chi^2_{\alpha}$ JOBB OLDALI $C_f=\chi^2_{1-\alpha}$	JOBB OLDALI $C_f=\chi^2_{1-\alpha}$		
Két mintás	Aszimptotikus Z-próba	Z-próba		Homogenitás-vizsgálat: χ^2-próba		NAGY MINTA, BÁRMILYEN ELOSZLÁSÚ SOKASÁGRA
	Z-próba t-próba		F-próba			CSAK NORMÁLIS ELOSZLÁSÚ SOKASÁGRA
	KÉTOLDALI $C_a=-Z_{1-\frac{\alpha}{2}}$ $C_f=-Z_{1-\frac{\alpha}{2}}$ BAL OLDALI $C_a=-Z_{1-\alpha}$ JOBB OLDALI $C_f=Z_{1-\alpha}$		KÉTOLDALI $C_a=\frac{1}{F_{1-\frac{\alpha}{2}}(\nu_2;\nu_1)}$ $C_f=F_{1-\frac{\alpha}{2}}(\nu_1;\nu_2)$ BAL OLDALI $C_a=\frac{1}{F_{1-\alpha}(\nu_2;\nu_1)}$ JOBB OLDALI $C_f=F_{1-\alpha}(\nu_1;\nu_2)$	JOBB OLDALI $C_f=\chi^2_{1-\alpha}$		
Több mintás	Variancia-analízis		Bartlett			
	JOBB OLDALI $C_f=F_{1-\alpha}(\nu_1;\nu_2)$		JOBB OLDALI $C_f=\chi^2_{1-\alpha}$			

9. REGRESSZIÓSZÁMÍTÁS

9.1. A kétváltozós regresszió egyenletei

9.1.1. A három fő regressziós modell egyenlete

LINEÁRIS MODELL

$$\hat{y} = b_0 + b_1 x$$

$$b_1 = \frac{\sum dx \cdot dy}{\sum d^2 x}$$

$$b_0 = \bar{y} - \bar{x} \cdot b_1$$

HATVÁNYKITEVŐS MODELL

$$\hat{y} = b_0 \cdot x^{b_1}$$

$$b_1 = \frac{\sum d \lg x \cdot d \lg y}{\sum d^2 \lg x}$$

$$\lg b_0 = \overline{\lg y} - \overline{\lg x} \cdot b_1$$

EXPONENCIÁLIS MODELL

$$\hat{y} = b_0 \cdot b_1^{\,x}$$

$$\lg b_1 = \frac{\sum dx \cdot d \lg y}{\sum d^2 x}$$

$$\lg b_0 = \overline{\lg y} - \bar{x} \cdot \lg b_1$$

9.1.2. A reziduumok

$$e_i = y_i - \hat{y}_i$$

mateking.hu

9.1.3. A reziduális szórás

$$s_e^* = \sqrt{\frac{SSE}{n}} = \sqrt{\frac{\sum e_i^2}{n}} = \sqrt{\frac{\sum (y_i - \hat{y}_i)}{n}}$$

9.1.4. A lineáris korrelációs együttható

$$r = \frac{\sum dx \cdot dy}{\sqrt{\sum d^2 x \cdot \sum d^2 y}}$$

9.1.5. Elaszticitás

$$El(\hat{y}, x_j) = \frac{b_j x_j}{\hat{y}}$$

9.1.6. A modell ereje (%)

$$PRE = r^2 = R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

$$SST = SSR + SSE$$

$$SST = \sum d^2 y \qquad SSR = \sum \left(\hat{y}_i - \bar{\hat{y}}\right)^2 = \beta_1^2 \sum d^2 x \qquad SSE = \sum \left(y_i - \hat{y}_i\right)^2 = \sum e_i^2$$

9.2. Többváltozós lineáris regresszió

9.2.1. A többváltozós lineáris regresszió egyenlete

$$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_k x_k + \varepsilon$$

$$X = \begin{pmatrix} 1 & x_{11} & x_{21} & \ldots & x_{k1} \\ 1 & x_{12} & x_{22} & \ldots & x_{k2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_{1n} & x_{2n} & \ldots & x_{kn} \end{pmatrix} \qquad \underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}$$

$$\underline{\hat{\beta}} = \left(X^T X\right)^{-1} X^T \underline{y}$$

9.2.2. A reziduumok

$$e_i = y_i - \hat{y}_i$$

9.2.3. A reziduális szórás

$$s_e = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n-k-1}}$$

9.2.4. Az elaszticitás

$$El(\hat{y}, x_j) = \frac{\hat{\beta}_j x_j}{\hat{y}}$$

9.2.5. A korrelációs mátrix

$$R = \begin{pmatrix} 1 & & & & \\ r_{1y} & 1 & & & \\ r_{2y} & r_{21} & 1 & & \\ \ldots & \ldots & \ldots & \ldots & \\ r_{ky} & r_{k1} & r_{k2} & \ldots & 1 \end{pmatrix}$$

9.2.6. A parciális korrelációs együtthatók

$$r_{yi.} = \frac{-q_{yj}}{\sqrt{q_{yy} q_{jj}}} \quad \text{ahol} \quad R^{-1} = \begin{pmatrix} q_{yy} & q_{1y} & \ldots & q_{1k} \\ q_{y1} & q_{11} & \ldots & q_{1k} \\ \ldots & \ldots & \ldots & \ldots \\ q_{yk} & q_{1k} & \ldots & q_{kk} \end{pmatrix}$$

mateking.hu

9.3. A standard lineáris modell

STANDARD LINEÁRIS MODELL FELTÉTELEI:

I. A magyarázó változók nem valószínűségi változók.

II. A magyarázó változók lineárisan független rendszert alkotnak.

III. Az eredményváltozó közel lineáris függvénye a magyarázó változóknak.

IV. Az ε hibatag feltételes eloszlása normális, várható értéke nulla.

V. Az ε hibatag különböző x-ekhez tartozó értékei korrelálatlanok.

9.3.1. A paraméterek intervallumbecslése

Kétváltozós eset

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1$$

Paraméterek becslése

$$\hat{\beta}_i \pm t_{1-\frac{\alpha}{2}} \cdot (n-k-1) \cdot s_{\hat{\beta}_i}$$

Regresszió becslése

$$\hat{y}_* \pm t_{1-\frac{\alpha}{2}} \cdot (n-k-1) \cdot s_{\hat{y}_*}$$

n=megfigyelések száma
k=paraméterek száma

$$s_e = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n-k-1}}$$

$$s_{\hat{\beta}_0} = s_e \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum d^2 x}}$$

$$s_{\hat{\beta}_1} = \frac{s_e}{\sqrt{\sum d^2 x}}$$

Többváltozós eset

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_k x_k$$

Paraméterek becslése

$$\hat{\beta}_i \pm t_{1-\frac{\alpha}{2}} \cdot (n-k-1) \cdot s_{\hat{\beta}_i}$$

Regresszió becslése

$$\hat{y}_* \pm t_{1-\frac{\alpha}{2}} \cdot (n-k-1) \cdot s_{\hat{y}_*}$$

n=megfigyelések száma
k=paraméterek száma

$$s_e = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n-k-1}}$$

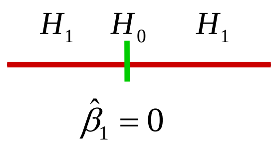
$$s_{\hat{\beta}_i} = s_e \cdot \sqrt{\left[\left(X^T \cdot X\right)^{-1}\right]_{ii}}$$

$$s_{\hat{y}_*} = s_e \cdot \sqrt{X_*^T \left(X^T X\right)^{-1} X_*}$$

9.3.2. A paraméterek tesztelése

t-próba:

$$t = \frac{\hat{\beta}_1 - \beta_1}{s_{\hat{\beta}_1}} \quad \text{ahol} \quad s_{\hat{\beta}_1} = \frac{s_e}{\sqrt{\sum d^2 x}}$$



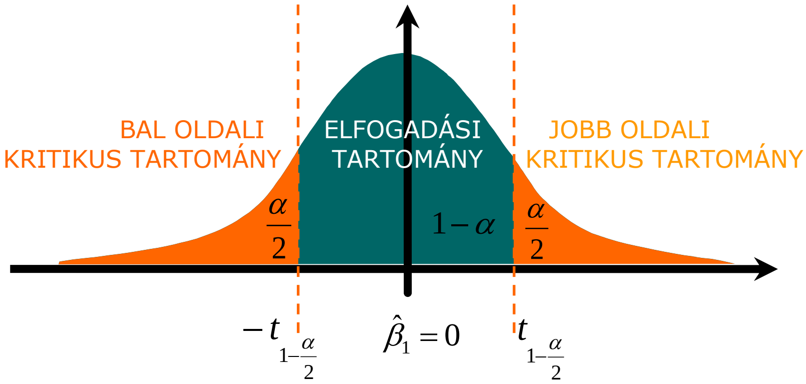
KÉTOLDALI KRITIKUS TARTOMÁNY

$H_0: \hat{\beta}_1 = 0$

$H_1: \hat{\beta}_1 \neq 0$

BAL OLDALI KRITIKUS ÉRTÉK: $-t_{1-\frac{\alpha}{2}}$

JOBB OLDALI KRITIKUS ÉRTÉK: $t_{1-\frac{\alpha}{2}}$



9.3.3. A modell egészének tesztelése

mateking.hu

A próbafüggvény

$$F(v_1; v_2) = \frac{SSR/k}{SSE/(n-k-1)}$$

$H_0: \hat{\beta}_i = 0$

$H_1: \hat{\beta}_i \neq 0$.

A két szabadságfok $v_1 = k$ és $v_2 = n-k-1$, a próba jobb oldali kritikus értékkel hajtandó végre:

JOBB OLDALI KRITIKUS ÉRTÉK: $F_{1-\alpha}(v_1; v_2)$

VARIANCIAANALÍZIS-TÁBLÁZAT

SZÓRÓDÁS OKA	NÉGYZETÖSSZEG	SZABADSÁG-FOK	ÁTLAGOS NÉGYZETÖSSZEG	F
Regresszió	SSR	k	$MSR = \frac{SSR}{k}$	$F = \frac{MSR}{MSE}$
Hiba	SSE	$n-k-1$	$MSE = \frac{SSE}{n-k-1}$	
Teljes	SST	$n-1$		

9.3.4. A multikollinearitás

$$VIF_j = \frac{1}{1-R_j^2}$$

A képletben szereplő R_j^2 a j-edik magyarázó változó és az összes többi magyarázó változó közti determinációs együttható.

9.3.5. Autokorreláció tesztelése

A próbafüggvény

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=2}^{n} e_t^2}$$

A szignifikanciaszint α, a próba elvégzése pedig az alábbi módon történik:
d_L és d_U értékeket kikeressük a táblázatból,
n=a megfigyelések száma,
k=a magyarázó változók száma
végül megnézzük a próbafüggvény melyik tartományba esik.