Pouzdani intervali

Pretpostavke	Pouzdani intervali za	Statistika	Razdioba
$X_1, \ldots, X_n \sim N(\mu, \sigma^2)$	μ (σ poznato)	$Z = \frac{\overline{X}_n - \mu}{\sigma}\sqrt{n}$	$\mathrm{N}(0,1)$
	μ (σ nepoznato)	$T = \frac{\overline{X}_n - \mu}{S_n}\sqrt{n}$	$\mathrm{t}(n-1)$
	σ^2 (μ nepoznato)	$V = \frac{n-1}{\sigma^2} S_n^2$	$\chi^2(n-1)$
	σ^2 (μ poznato)	$U = \sum_{n=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2$	$\chi^2(n)$
$X_1, \ldots, X_n$ sl. uzorak, $n > 30$, $\mu = EX_1, \sigma^2 = \mathrm{Var} X_1 < +\infty$		$Z = \frac{\overline{X}_n - \mu}{\sigma}\sqrt{n}$	$\mathrm{AN}(0,1)$

Linearni regresijski model $y = \alpha + \beta x$

Konstrukcija pouzdanih intervala za α i β:

$$\frac{\hat{\alpha} - \alpha}{\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\overline{x}^2}{S_{xx}}}} \sim t(n-2), \quad \frac{\hat{\beta} - \beta}{\hat{\sigma}\sqrt{\frac{1}{S_{xx}}}} \sim t(n-2),$$

gdje je

$$\hat{\sigma} = \sqrt{\frac{SSE}{n-2}}, \quad SSE = S_{yy} - \hat{\beta}^2 S_{xx}.$$

Konstrukcija pouzdanih intervala za σ^2:

$$(n-2)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-2)$$

Konstrukcija pouzdanih intervala za $E[Y|x = x_0]$:

$$\sqrt{\frac{n-2}{SSE}} \cdot \frac{\hat{\alpha} + \hat{\beta}x - (\alpha + \beta x)}{\sqrt{\frac{1}{n} + \frac{(x - \overline{x})^2}{S_{xx}}}} \sim t(n-2)$$

Konstrukcija pouzdanih intervala za Y u $x = x_0$:

$$\sqrt{\frac{n-2}{SSE}} \cdot \frac{\hat{\alpha} + \hat{\beta}x - Y}{\sqrt{1 + \frac{1}{n} + \frac{(x - \overline{x})^2}{S_{xx}}}} \sim t(n-2)$$

Statistički testovi

Pretpostavke		H_0	H_1	Testna statistika	Razdioba
$X_1, \ldots, X_n \sim \mathrm{N}(\mu, \sigma^2)$	σ poznato	$\mu = \mu_0$	$\mu \neq \mu_0$	$Z = \frac{\overline{X}_n - \mu_0}{\sigma}\sqrt{n}$	$\mathrm{N}(0,1)$
			$\mu < \mu_0$		
			$\mu > \mu_0$		
	σ nepoznato		$\mu \neq \mu_0$	$T = \frac{\overline{X}_n - \mu_0}{S_n}\sqrt{n}$	$\mathrm{t}(n-1)$
			$\mu < \mu_0$		
			$\mu > \mu_0$		
	μ nepoznato	$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$V = \frac{n-1}{\sigma_0^2} S_n^2$	$\chi^2(n-1)$
			$\sigma^2 < \sigma_0^2$		
			$\sigma^2 > \sigma_0^2$		
$X_1, \ldots, X_n$ sl. uzorak $n > 30, \mu = \mathrm{E}X_1$, $\sigma^2 = \mathrm{Var}X_1 < +\infty$		$\mu = \mu_0$	$\mu \neq \mu_0$	$Z = \frac{\overline{X}_n - \mu_0}{\sigma}\sqrt{n}$	$\mathrm{AN}(0,1)$
			$\mu < \mu_0$		
			$\mu > \mu_0$		
$X_1, \ldots, X_{n_1} \sim N(\mu_1, \sigma_1^2)$ $Y_1, \ldots, Y_{n_2} \sim N(\mu_2, \sigma_2^2)$ uzorci nezavisni	$\sigma_1^2 = \sigma_2^2$ nepoznato	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$T = \frac{\overline{X}_1 - \overline{X}_2}{S_d} \frac{1}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$\mathrm{t}(n_1 + n_2 - 2)$
			$\mu_1 < \mu_2$		
			$\mu_1 > \mu_2$	$S_d = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}}$	
	μ_1, μ_2, nepoznati	$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$	$F = \frac{S_1^2}{S_2^2}$	$\mathrm{F}(n_1 - 1, n_2 - 1)$
			$\sigma_1^2 < \sigma_2^2$		
			$\sigma_1^2 > \sigma_2^2$		

Test proporcija

Pretpostavke: $X_1, \ldots, X_{n_1}$ iz Bernoullijevog modela s vjerojatnosti uspjeha p_1
$Y_1, \ldots, Y_{n_2}$ iz Bernoullijevog modela s vjerojatnosti uspjeha p_2
uzorci nezavisni

$H_0\colon p_1 = p_2$

$H_1\colon p_1 \neq p_2$ ili $p_1 < p_2$ ili $p_1 > p_2$

$\hat{p}_1$= procjenitelj za p_1, $\hat{p}_2$= procjenitelj za p_2

Testna statistika: $T = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})}} \frac{1}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \overset{H_0}{\sim} \mathrm{AN}(0,1), \quad \hat{p} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}$

Jednofaktorska analiza varijance (ANOVA)

Pretpostavke: $X_{11}, X_{12}, \ldots, X_{1n_1} \sim N(\mu_1, \sigma^2)$
$X_{21}, X_{22}, \ldots, X_{2n_2} \sim N(\mu_2, \sigma^2)$
$\vdots$
$X_{k1}, X_{k2}, \ldots, X_{kn_k} \sim N(\mu_k, \sigma^2)$
uzorci nezavisni

$H_0\colon \mu_1 = \mu_2 = \cdots = \mu_k$
$H_1\colon$ ne H_0

$$n = \sum_{i=1}^{k} n_i, \quad \overline{X} = \frac{1}{n}\sum_{i=1}^{k} n_i \overline{X}_{i\cdot}, \quad SST = \sum_{i=1}^{k} n_i(\overline{X}_{i\cdot} - \overline{X})^2, \quad SSE = \sum_{i=1}^{k}(n_i - 1)S_{i\cdot}^2$$

Testna statistika: $F = \frac{MST}{MSE} \overset{H_0}{\sim} F(k-1, n-k), \quad MST = \frac{SST}{k-1}, \quad MSE = \frac{SSE}{n-k}$
Kritično područje: $[f_\alpha(k-1, n-k), +\infty\rangle$

χ^2-test o pripadnosti distribuciji

Opažene frekvencije: $N_i, i = 1, \ldots, k$
Očekivane frekvencije: $n_i, i = 1, \ldots, k$
Testna statistika: $H = \sum_{i=1}^{k} \frac{(N_i - n_i)^2}{n_i}$
Kritično područje: $[\chi^2_\alpha(k-d-1), +\infty\rangle$, d-broj procijenjenih parametara

χ^2-test o nezavisnosti

$X\backslash Y$	b_1	b_2	$\cdots$	b_c	Σ
a_1	N_{11}	N_{12}	$\cdots$	N_{1c}	N_1
a_2	N_{21}	N_{22}	$\cdots$	N_{2c}	N_2
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
a_r	N_{r1}	N_{r2}	$\cdots$	N_{rc}	N_r
Σ	M_1	M_2	$\cdots$	M_c	n

Testna statistika: $H_n = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(N_{ij} - n\hat{p}_i\hat{q}_j)^2}{n\hat{p}_i\hat{q}_j}$

$$\hat{p}_i = \frac{N_i}{n}, \hat{q}_j = \frac{M_j}{n}$$

Kritično područje: $[\chi^2_\alpha((r-1)(c-1)), +\infty\rangle$

χ^2-test o homogenosti

populacija$\backslash X$	a_1	a_2	$\cdots$	a_k	Σ
1	N_{11}	N_{12}	$\cdots$	N_{1k}	n_1
2	N_{21}	N_{22}	$\cdots$	N_{2k}	n_2
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
m	N_{m1}	N_{m2}	$\cdots$	N_{mk}	n_m
Σ	M_1	M_2	$\cdots$	M_k	n

Testna statistika: $H_n = \sum_{i=1}^{m}\sum_{j=1}^{k} \frac{(N_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$

$$\hat{n}_{ij} = \frac{n_i M_j}{n}$$

Kritično područje: $[\chi^2_\alpha((m-1)(k-1)), +\infty\rangle$

Kolmogorov Smirnovljev test

H_0: $F=F_0$, (F_0 - konkretna neprekidna distribucija)
H_1: ne H_0
Testna statistika: $D_n = \max_{1\le i\le n}\{\max\{|\frac{i-1}{n} - F_0(x_{(i)})|, |\frac{i}{n} - F_0(x_{(i)})|\}\}$
Kritično područje: $[d_\alpha(n), +\infty\rangle$