ESTYMACJA
I
WERYFIKACJA HIPOTEZ

Materiały pomocnicze do wykładu i ćwiczeń,
dr hab. Ewa Frątczak, SGH.

ESTYMACJA PRZEDZIAŁOWA

1. Przedział ufności dla średniej *m* w populacji normalnej ze znanym odchyleniem standardowym

$$P\left(\bar{x} - u_\alpha \cdot \frac{\sigma}{\sqrt{n}} < m < \bar{x} + u_\alpha \cdot \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

2. Przedział ufności dla średniej *m* w populacji normalnej z nieznanym odchyleniem standardowym

$$P\left(\bar{x} - t_{\alpha,n-1} \cdot \frac{S(x)}{\sqrt{n-1}} < m < \bar{x} + t_{\alpha,n-1} \cdot \frac{S(x)}{\sqrt{n-1}}\right) = 1 - \alpha$$

$$P\left(\bar{x} - t_{\alpha,n-1} \cdot \frac{s}{\sqrt{n}} < m < \bar{x} + t_{\alpha,n-1} \cdot \frac{s}{\sqrt{n}}\right) = 1 - \alpha$$

3. Przedział ufności dla średniej *m* w populacji o nieznanym rozkładzie, gdzie liczebność próby jest duża (*n*>120) $\bar{X}$ ma rozkład graniczny $N\left(m; \frac{\sigma}{\sqrt{n}}\right)$

$$P\left(\bar{x} - u_\alpha \cdot \frac{\sigma}{\sqrt{n}} < m < \bar{x} + u_\alpha \cdot \frac{\sigma}{\sqrt{n}}\right) \equiv 1 - \alpha$$

4. Przedział ufności dla wariancji σ^2 w populacji normalnej

$$P\left(\frac{nS^2}{\chi^2_{\alpha/2,n-1}} < \sigma^2 < \frac{nS^2}{\chi^2_{1-\alpha/2,n-1}}\right) \equiv 1 - \alpha$$

$$P\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}\right) \equiv 1 - \alpha$$

5. Przedział ufności dla odchylenia standardowego (duża próba)

$$\mathrm{E}(S(X)) = \mathrm{D}(X)$$

$$\mathrm{D}(S(X)) = \frac{\mathrm{D}(X)}{\sqrt{2n}} = \frac{\mathrm{S}(X)}{\sqrt{2n}}$$

$$P\left(S(X) - u_\alpha \cdot \frac{S(X)}{\sqrt{2n}} < \mathrm{D}(X) < S(X) + u_\alpha \cdot \frac{S(X)}{\sqrt{2n}} \right) \equiv 1 - \alpha$$

$$P\left(s - u_\alpha \cdot \frac{s}{\sqrt{2n}} < \mathrm{D}(X) < s + u_\alpha \cdot \frac{s}{\sqrt{2n}} \right) \equiv 1 - \alpha$$

6. Przedział ufności dla parametru p w rozkładzie dwumianowym

$$P\left(w_i - u_\alpha \cdot \sqrt{\frac{w_i(1-w_i)}{n}} < p < w_i + u_\alpha \cdot \sqrt{\frac{w_i(1-w_i)}{n}} \right) \equiv 1 - \alpha \ , \qquad w_i = \frac{n_i}{n}$$

$$P\left(\hat{p} - u_\alpha \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + u_\alpha \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right) \equiv 1 - \alpha$$

7. Minimalna liczebność próby

$$u_\alpha \cdot \frac{\sigma}{\sqrt{n}} \le d \Rightarrow n \ge \frac{u^2{}_\alpha \sigma^2}{d^2}$$

$$u_\alpha \cdot \sqrt{\frac{p(1-p)}{n}} \le d \Rightarrow n \ge \frac{u^2{}_\alpha\, p(1-p)}{d^2}$$

$$n \ge \frac{u_\alpha^2}{4d^2}$$

PARAMETRYCZNE TESTY ISTOTNOŚCI

Hipoteza	Statystyka	Rozkład	obszar krytyczny

1. Populacja generalna $N(m;\sigma)$, σ - znane

$H_0: m = m_0$ $H_1: m \neq m_0$	$U = \frac{\overline{X} - m_0}{\sigma}\sqrt{n}$	$N(0;1)$	$\lvert U \rvert \geq u_\alpha$

2. Populacja generalna $N(m;\sigma)$, σ - nieznane

$H_0: m = m_0$ $H_1: m \neq m_0$	$t = \frac{\overline{X} - m_0}{S}\sqrt{n-1}$	t-Studenta o n-1 st. swobody	$\lvert t \rvert \geq t_{\alpha,n-1}$
	$t = \frac{\overline{X} - m_0}{S}\sqrt{n}$		

3. Populacja generalna ma rozkład dowolny z nieznanymi parametrami

$H_0: m = m_0$ $H_1: m \neq m_0$	$U = \frac{\overline{X} - m_0}{\sigma}\sqrt{n}$	asymptot. $N(0;1)$	$\lvert U \rvert \geq u_\alpha$
	$U = \frac{\overline{X} - m_0}{s}\sqrt{n}$		

4. Dwie populacje normalne o znanych wariancjach, n_1, n_2 - liczebności prób

$H_0: m_1 = m_2$ $H_1: m_1 \neq m_2$	$U = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$N(0;1)$	$\lvert U \rvert \geq u_\alpha$

5. Dwie populacje normalne o nieznanej, ale jednakowej wariancji

n_1, n_2 - liczebności prób

$H_0: m_1 = m_2$	$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{n_1 S_1^2 + n_2 S_2^2}} \cdot \sqrt{\frac{n_1 n_2}{n_1 + n_2}(n_1 + n_2 - 2)}$		$\lvert t \rvert \geq t_{\alpha, n_1+n_2-2}$

$H_1 : m_1 \neq m_2$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$ gdzie s_p^2 to wariancja z połączonych prób

Statystyka t posiada rozkład t-Studenta o $v = n_1 + n_2 - 2$ stopniach swobody

6. Dwie populacje o dowolnych rozkładach, n_1, n_2 - liczebności prób

$H_0 : m_1 = m_2$ $$U = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$ asymptot. $N(0;1)$ $|U| \geq u_\alpha$

$H_1 : m_1 \neq m_2$

7. Próbę losową stanowią uporządkowane pary zmiennych losowych X_{i1} i X_{i2}.

$H_0 : m_R = m_0$ $$t = \frac{\overline{R} - m_0}{S_R}\sqrt{n-1}$$ t-Studenta $|t| \geq t_{\alpha, n-1}$

$H_1 : m_R \neq m_0$ o *n*-1 st. swobody

$$t = \frac{\overline{R} - m_0}{s_R}\sqrt{n}$$

Różnice zmiennych $R_i = X_{i1} - X_{i2}$ mają w populacji rozkład $N(m_R; \sigma_R)$

gdzie $\overline{R} = \frac{1}{n}\sum_{i=1}^{n} R_i = \frac{1}{n}\sum_{i=1}^{n}(X_{i1} - X_{i2})$, $S_R = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_{i1} - X_{i2})}$

$$s_R = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_{i1} - X_{i2})}$$

8. Populacja generalna ma rozkład dwupunktowy z parametrem *p*.

$H_0 : p = p_0$ $$U = \frac{W - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$ asymptot. $N(0;1)$ $|U| \geq u_\alpha$

$H_1 : p \neq p_0$

$$U = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Statystyka $W = \frac{X}{n}$ ma rozkład normalny $N\left(p; \sqrt{\frac{p(1-p)}{n}}\right)$

9. Dwie populacje generalne o rozkładzie dwupunktowym $W_1 = \frac{X_1}{n}$, $W_2 = \frac{X_2}{n}$,

$H_0 : p_1 = p_2$ $\qquad U = \frac{W_1 - W_2}{\sqrt{\frac{\tilde{p}(1-\tilde{p})}{n}}}$ $\qquad$ asymptot. $N(0;1)$ $\qquad |U| \geq u_\alpha$

$H_1 : p_1 \neq p_2$

$$U = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$n = \frac{n_1 n_2}{n_1 + n_2}$, $\tilde{p} = \frac{m_1 + m_2}{n_1 + n_2}$ $\qquad \hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$

$W_1 = W_2$ ma rozkład asymptotyczny normalny $N\left(p_1 - p_2; \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}\right)$

Uwaga. Hipoteza H_1 może być zapisana jako:

$H_0 : Q = Q_0$ $\qquad$ lub $\qquad H_0 : Q = Q_0$

$H_1 : Q > Q_0$ $\qquad\qquad\quad H_1 : Q < Q_0$

Obszary odrzuceń H_0 możemy określać dla:

$U \geq u_{2\alpha}$, $\qquad U \leq u_{2\alpha}$

$t \geq t_{2\alpha,v}$ $\qquad t \leq t_{2\alpha,v}$

v - liczba stopni swobody

10. Populacja ma rozkład normalny o nieznanych parametrach

$H_0 : \sigma^2 = \sigma_0^2$ $\qquad \chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i - \bar{x})^2$ $\qquad \chi^2$ o n-1 st.swob. $\qquad \chi^2 \geq \chi^2_{\alpha,n-1}$

$H_1 : \sigma^2 \neq \sigma_0^2$

11. Dwie populacje o rozkładach normalnych

$H_0 : \sigma_1^2 = \sigma_2^2$ $\qquad \chi^2 = \frac{n_1\tilde{s}_1^2/(n_1-1)}{n_2\tilde{s}_2^2/(n_2-1)}$ $\qquad$ F z $\begin{matrix} v_1 = n_1 - 1 \\ v_2 = n_2 - 1 \end{matrix}$ st.swob. $\qquad F \geq F_{\frac{\alpha}{2}}$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

NIEPARAMETRYCZNE TESTY ISTOTNOŚCI

Hipoteza	**Statystyka**	**Rozkład**	**obszar krytyczny**

1. Rozkład populacji dowolny, duża próba losowa

$H_0 : F(x) = F_0(x)$ $H_0 : F(x) \neq F_0(x)$	$\chi^2 = \sum \frac{\left(n_i - \hat{n}_i\right)^2}{\hat{n}_i}$	asymptot. rozkład χ^2 o v=r+k-1 stopniach swobody	$\chi^2 \geq \chi^2_\alpha$

$\hat{n}_i = n \cdot p_i$

2. Zmienna losowa w populacji ma rozkład ciągły określony dystrybuantą $F_0(x)$

$H_0 : F(x) = F_0(x)$ $H_0 : F(x) \neq F_0(x)$	$\lambda = D\sqrt{n}$	asymptot. rozkład λ Kołmogorowa	$\lambda \geq \lambda_\alpha$

$D = \sup_x \left| F_n(x) - F_0(x) \right|$

3. Dwie zmienne w populacji mają rozkład normalny określony dystrybuantami $F_1(x), F_2(x)$

$H_0 : F_1(x) = F_2(x)$ $H_0 : F_1(x) \neq F_2(x)$	$\lambda = D^* \sqrt{n}$ $n = \frac{n_1 n_2}{n_1 + n_2}$	asymptot. rozkład λ Kołmogorowa	$\lambda \geq \lambda_\alpha$

$D^* = \sup_x \left| F_{n_1}(x) - F_{n_2}(x) \right|$