

**Podstawowe statystyki próbkkowe i funkcje w R wyliczające te statystyki**

średnia z próby: $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$ , $> \text{mean}(\text{dane})$	wariancja z próby: $s^2 = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2$ $> \text{var}(\text{dane})$
---	--

**Oznaczenia podstawowych kwantyli i funkcje w R wyliczające te kwantyle**

dla rozkładu normalnego $u_\alpha$ $> \text{qnorm}(\alpha)$	dla rozkładu t-Studenta $t_{\alpha,n}$ $> \text{qt}(\alpha, n)$	dla rozkładu chi-kwadrat $\chi_{\alpha,n}^2$ $> \text{qchisq}(\alpha, n)$
---	---	---

**UWAGA:** Jeśli wyznaczone wartości statystyk testowych należą do odpowiednich zbiorów krytycznych, to  $H_0$  odrzucamy.

**Weryfikacje hipotez dotyczących wartości średniej na poziomie istotności  $\alpha$** 

**Model I.**  $X \sim N(\mu, \sigma^2)$ ,  $\mu$  - nieznane,  $\sigma$  - znane.

Hipoteza zerowa  $H_0 : \mu = \mu_0$ . Statystyka testowa  $U = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$ .

Hipoteza alternatywna $H_1 : \mu \neq \mu_0$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha/2}) \cup (u_{1-\alpha/2}; +\infty)$	Hipoteza alternatywna $H_1 : \mu > \mu_0$ Zbiór krytyczny $W = (u_{1-\alpha}; +\infty)$	Hipoteza alternatywna $H_1 : \mu < \mu_0$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha})$
---	--	---

**Model II (t.test).**  $X \sim N(\mu, \sigma^2)$ ,  $\mu$  - nieznane,  $\sigma$  - nieznane.

Hipoteza zerowa  $H_0 : \mu = \mu_0$ . Statystyka testowa  $T = \frac{\bar{X} - \mu_0}{s} \sqrt{n}$ .

Hipoteza alternatywna $H_1 : \mu \neq \mu_0$ Zbiór krytyczny $W = (-\infty; -t_{1-\alpha/2,n-1}) \cup (t_{1-\alpha/2,n-1}; +\infty)$	Hipoteza alternatywna $H_1 : \mu > \mu_0$ Zbiór krytyczny $W = (t_{1-\alpha,n-1}; +\infty)$	Hipoteza alternatywna $H_1 : \mu < \mu_0$ Zbiór krytyczny $W = (-\infty; -t_{1-\alpha,n-1})$
---	--	---

**Model III.**  $X$  ma rozkład dowolny (próba duża:  $n \geq 100$ ).

Hipoteza zerowa  $H_0 : \mu = \mu_0$ . Statystyka testowa  $U = \frac{\bar{X} - \mu_0}{s} \sqrt{n}$ .

Hipoteza alternatywna $H_1 : \mu \neq \mu_0$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha/2}) \cup (u_{1-\alpha/2}; +\infty)$	Hipoteza alternatywna $H_1 : \mu > \mu_0$ Zbiór krytyczny $W = (u_{1-\alpha}; +\infty)$	Hipoteza alternatywna $H_1 : \mu < \mu_0$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha})$
---	--	---

**Model IV (prop.test).**  $X$  ma rozkład dwupunktowy  $P(X = 1) = p$ ,  $P(X = 0) = q = 1 - p$ ,  $p$  - nieznane,  $n\hat{p} \geq 5$  i  $n\hat{q} \geq 5$ , gdzie  $\hat{p} = \frac{k}{n} = \frac{\text{ilość sukcesów}}{\text{ilość prób}}$ ,  $\hat{q} = 1 - \hat{p}$ .

Hipoteza zerowa  $H_0 : p = p_0$ . Statystyka testowa  $U = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}.$

Hipoteza alternatywna $H_1 : p \neq p_0$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha/2}) \cup (u_{1-\alpha/2}; +\infty)$	Hipoteza alternatywna $H_1 : p > p_0$ Zbiór krytyczny $W = (u_{1-\alpha}; +\infty)$	Hipoteza alternatywna $H_1 : p < p_0$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha})$
---	--	---

Jeśli w **modelu IV** nie jest spełnione założenie, że  $n\hat{p} \geq 5$  i  $n\hat{q} \geq 5$ , to zamiast prop.test używamy testu dokładnego **binom.test**.

**Weryfikacja hipotezy dotyczącej jednej wariancji na poziomie istotności  $\alpha$** 

**Model.**  $X \sim N(\mu, \sigma^2)$ ,  $\mu$  - nieznane,  $\sigma$  - nieznane.

Hipoteza zerowa  $H_0 : \sigma^2 = \sigma_0^2$ . Statystyka testowa  $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$ .

Hipoteza alternatywna $H_1 : \sigma^2 \neq \sigma_0^2$ Zbiór krytyczny $W = (0, \chi_{\alpha/2;n-1}^2) \cup (\chi_{1-\alpha/2;n-1}^2, +\infty)$	Hipoteza alternatywna $H_1 : \sigma^2 > \sigma_0^2$ Zbiór krytyczny $W = (\chi_{1-\alpha;n-1}^2, +\infty)$	Hipoteza alternatywna $H_1 : \sigma^2 < \sigma_0^2$ Zbiór krytyczny $W = (0; \chi_{\alpha;n-1}^2)$
--	---	---

### Weryfikacje hipotez dotyczących dwóch średnich na poziomie istotności $\alpha$

**UWAGA:** jeżeli wyznaczone wartości statystyk ( $U$  lub  $T$ ) należą do odpowiednich zbiorów krytycznych, to  $H_0$  odrzucamy.

**Model I.**  $X \sim N(\mu_1, \sigma_1^2)$ ,  $Y \sim N(\mu_2, \sigma_2^2)$ ,  $\mu_1, \mu_2$  - nieznane,  $\sigma_1, \sigma_2$  - znane; dysponujemy niezależnymi próbami losowymi z tych populacji.

Hipoteza zerowa  $H_0 : \mu_1 = \mu_2$ . Statystyka testowa  $U = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ .

Hipoteza alternatywna	$H_1 : \mu_1 \neq \mu_2$	Hipoteza alternatywna	$H_1 : \mu_1 > \mu_2$	Hipoteza alternatywna	$H_1 : \mu_1 < \mu_2$
Zbior krytyczny $W = (-\infty; -u_{1-\alpha/2}) \cup \langle u_{1-\alpha/2}; +\infty)$	Zbior krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Zbior krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Zbior krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Zbior krytyczny $W = (-\infty; -u_{1-\alpha})$	Zbior krytyczny $W = (-\infty; -u_{1-\alpha})$

<b>Model II.(unpaired t-test: t.test(...,paired=FALSE, var.equal=TRUE))</b> $X \sim N(\mu_1, \sigma_1^2)$ , $Y \sim N(\mu_2, \sigma_2^2)$ , $\mu_1, \mu_2$ - nieznane, $\sigma_1, \sigma_2$ - nieznane, ale takie, że $\sigma_1 = \sigma_2$ ; dysponujemy niezależnymi próbami losowymi z tych populacji. Hipoteza zerowa $H_0 : \mu_1 = \mu_2$ . Statystyka testowa $T = \frac{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$ .					
Hipoteza alternatywna $H_1 : \mu_1 \neq \mu_2$	Hipoteza alternatywna $H_1 : \mu_1 > \mu_2$	Hipoteza alternatywna $H_1 : \mu_1 > \mu_2$	Hipoteza alternatywna $H_1 : \mu_1 > \mu_2$	Hipoteza alternatywna $H_1 : \mu_1 < \mu_2$	Hipoteza alternatywna $H_1 : \mu_1 < \mu_2$
Zbior krytyczny $W = (-\infty, -t_{1-\alpha/2, n_1+n_2-2}) \cup \langle t_{1-\alpha/2, n_1+n_2-2}; +\infty)$	Zbior krytyczny $W = \langle t_{1-\alpha, n_1+n_2-2}; +\infty)$	Zbior krytyczny $W = \langle t_{1-\alpha, n_1+n_2-2}; +\infty)$	Zbior krytyczny $W = \langle t_{1-\alpha, n_1+n_2-2}; +\infty)$	Zbior krytyczny $W = (-\infty, -t_{1-\alpha, n-1})$	Zbior krytyczny $W = (-\infty, -t_{1-\alpha, n-1})$

Jesli w modelu II nie jest spełnione założenie, że  $\sigma_1 = \sigma_2$ , to zamiast t.test(...,paired=FALSE, var.equal=TRUE) należy użyć t.test(...,paired=TRUE).

<b>Model III.(paired t-test: t.test(..., paired=TRUE))</b> $X - Y \sim N(\mu_Z, \sigma_Z^2)$ , $\mu_Z, \sigma_Z$ - nieznane; dysponujemy parami obserwacji, gdzie pary są wzajemnie niezależne. Hipoteza zerowa $H_0 : \mu_1 = \mu_2$ . Statystyka testowa $T = \frac{\bar{z}}{s_z} \sqrt{n}$ , gdzie $z_i = x_i - y_i, i = 1, 2, \dots, n$ .					
Hipoteza alternatywna $H_1 : \mu_1 \neq \mu_2$	Hipoteza alternatywna $H_1 : \mu_1 > \mu_2$	Hipoteza alternatywna $H_1 : \mu_1 > \mu_2$	Hipoteza alternatywna $H_1 : \mu_1 > \mu_2$	Hipoteza alternatywna $H_1 : \mu_1 < \mu_2$	Hipoteza alternatywna $H_1 : \mu_1 < \mu_2$
Zbior krytyczny $W = (-\infty; -t_{1-\alpha/2, n-1}) \cup \langle t_{1-\alpha/2, n-1}; +\infty)$	Zbior krytyczny $W = \langle t_{1-\alpha, n-1}; +\infty)$	Zbior krytyczny $W = \langle t_{1-\alpha, n-1}; +\infty)$	Zbior krytyczny $W = (-\infty; -t_{1-\alpha, n-1})$	Zbior krytyczny $W = (-\infty; -t_{1-\alpha, n-1})$	Zbior krytyczny $W = (-\infty; -t_{1-\alpha, n-1})$

**Model IV.** Cechy  $X, Y$  mają rozkłady dowolne ( $n_1 \geq 100, n_2 \geq 100$ ),  $\mu_1, \mu_2, \sigma_1, \sigma_2$  - nieznane; dysponujemy niezależnymi próbami losowymi z tych populacji.

Hipoteza alternatywna $H_1 : p_1 \neq p_2$	Hipoteza alternatywna $H_1 : p_1 > p_2$	Hipoteza alternatywna $H_1 : p_1 > p_2$	Hipoteza alternatywna $H_1 : p_1 > p_2$	Hipoteza alternatywna $H_1 : p_1 < p_2$	Hipoteza alternatywna $H_1 : p_1 < p_2$
Zbior krytyczny $W = (-\infty; -u_{1-\alpha/2}) \cup \langle u_{1-\alpha/2}; +\infty)$	Zbior krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Zbior krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Zbior krytyczny $W = (-\infty; -u_{1-\alpha})$	Zbior krytyczny $W = (-\infty; -u_{1-\alpha})$	Zbior krytyczny $W = (-\infty; -u_{1-\alpha})$

<b>Model V (prop.test).</b> Cechy $X, Y$ mają rozkłady dwupunktowe, $P(X = 1) = p_1 = 1 - P(X = 0)$ , $P(Y = 1) = p_2 = 1 - P(Y = 0)$ , $p_1, p_2$ - nieznane, $n_1 \hat{p}_1 \geq 5$ i $n_2 \hat{p}_2 \geq 5$ i $n_2(1 - \hat{p}_2) \geq 5$ .					
Hipoteza alternatywna $H_1 : p_1 \neq p_2$	Hipoteza alternatywna $H_1 : p_1 > p_2$	Hipoteza alternatywna $H_1 : p_1 > p_2$	Hipoteza alternatywna $H_1 : p_1 > p_2$	Hipoteza alternatywna $H_1 : p_1 < p_2$	Hipoteza alternatywna $H_1 : p_1 < p_2$
Zbior krytyczny $W = (-\infty; -u_{1-\alpha/2}) \cup \langle u_{1-\alpha/2}; +\infty)$	Zbior krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Zbior krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Zbior krytyczny $W = (-\infty; -u_{1-\alpha})$	Zbior krytyczny $W = (-\infty; -u_{1-\alpha})$	Zbior krytyczny $W = (-\infty; -u_{1-\alpha})$

Jesli w modelu V nie jest spełnione założenie, że  $n_1, n_2$  są wystarczająco duże, to zamiast prop.test należy zastosować dokładny test Fishera fisher.test oparty na rozkładzie hipergeometrycznym.

<b>Weryfikacja hipotezy dotyczącej równości dwóch wariancji na poziomie istotności <math>\alpha</math></b>	
<b>UWAGA:</b> jeżeli wyznaczona wartość statystyki $F$ należy do $W$ , to $H_0$ odrzucamy.	
<b>Model I. (test F: var.test)</b>	
$X \sim N(\mu_1, \sigma_1^2)$ , $Y \sim N(\mu_2, \sigma_2^2)$ , $\mu_1, \mu_2, \sigma_1, \sigma_2$ - nieznane, dysponujemy niezależnymi próbami losowymi z tych populacji.	
Hipoteza zerowa $H_0 : \sigma_1^2 = \sigma_2^2$ . Statystyka testowa $F = s_1^2/s_2^2$ (w liczniku jest większa z wariancji).	
Hipoteza alternatywna $H_1 : \sigma_1^2 \neq \sigma_2^2$	Hipoteza alternatywna $H_1 : \sigma_1^2 > \sigma_2^2$
Zbiór krytyczny	Zbiór krytyczny
$W = \langle F(1 - \alpha/2, n_1 - 1, n_2 - 2); +\infty \rangle$	$W = \langle F(1 - \alpha, n_1 - 1, n_2 - 2); +\infty \rangle$
Powyżej $F(\alpha, n, m)$ oznacza kwantyl rozkładu F-Snedecora: $< \text{qf}(\alpha, n, m)$	

<b>Test zgodności <math>\chi^2</math>-Pearsona na poziomie istotności <math>\alpha</math></b>	
$H_0$ :	badana próba losowa pochodzi z zadanego rozkładu (lub rodziny rozkładów)
$H_1$ :	badana próba losowa nie pochodzi z zadanego rozkładu (lub rodziny rozkładów)
Statystyka testowa $\chi^2 = \sum_{j=1}^k \frac{(n_j - np_j^0)^2}{np_j^0}$ , gdzie $k$ - liczba klas,	
$n_j$ - liczba obserwacji, które znalazły się w $j$ -tej klasie, $n$ - liczność próby,	
$p_j^0$ - prawdopodobieństwo wpadnięcia obserwacji do $j$ -tej klasy przy założeniu prawdziwości $H_0$ (jeśli $H_0$ nie jest hipotezą prostą, to brakujące parametry rozkładu z $H_0$ wyznaczamy metodą NW)	
Zbiór krytyczny	
$W = \left\langle \chi_{1-\alpha, k-1-r}^2; +\infty \right\rangle$ , $r$ jest ilością parametrów szacowanych z próby	
<b>UWAGA:</b> jeżeli $\chi^2 \in W$ to hipotezę zerową $H_0$ odrzucamy.	

Test zgodności  $\chi^2$ -Pearsona jest zaimplementowany w R, niestety jedynie dla prostych hipotez  $H_0$ :

> `chisq.test(x, p)`

gdzie

- $x$  to wektor z licznosciami poszczególnych klas,
- $p$  to wektor z prawdopodobieństwami teoretycznymi  $p_j^0$  poszczególnych klas.