Wykłady 3-5: Testy statystyczne

Definicja. *Hipotezą statystyczną* nazywamy przypuszczenie dotyczące nieznanego rozkładu badanej cechy populacji, o prawdziwości lub fałszywości którego wnioskuje się na podstawie pobranej próby.

Przykład 3.1.

1. *Wysuwamy hipotezę, że badana cecha ma rozkład normalny.*
2. *Wiemy, że badana cecha ma rozkład normalny o nieznanej wartości średniej μ i znanym odchyleniu standardowym $\sigma = 1$. Wysuwamy hipotezę, że $\mu = 5$.*
3. *Dane są dwa zbiory obserwacji, np. wysokości plonów uzyskane podczas naważenia nawozem A i nawozem B.*
   (a) *Wysuwamy hipotezę, że oba zbiory można traktować jako pochodzące z populacji o tym samym rozkładzie.*
   (b) *Z wcześniejszych badań wiemy, że zbiory te można traktować jako pochodzące z populacji o rozkładach normalnych odpowiednio $\mathcal{N}(\mu_1, \sigma_1^2)$ i $\mathcal{N}(\mu_2, \sigma_2^2)$, gdzie parametry μ_1, μ_2, σ_1 i σ_2 są nieznane, ale takie, że $\sigma_1 = \sigma_2$. Wysuwamy przypuszczenie, że średnia wartość plonów przy nawożeniu nawozem A jest większa niż średnia wartość plonów przy nawożeniu nawozem B (tzn. że $\mu_1 > \mu_2$).*

Definicja. Hipotezy, które dotyczą wyłącznie wartości parametru lub parametrów rozkładu badanej cechy nazywamy *hipotezami parametrycznymi.* Hipotezy, które nie są hipotezami parametrycznymi nazywamy *hipotezami nieparametrycznymi.*

W przykładzie 3.1 hipotezy 2 i 3(b) są parametryczne, natomiast hipotezy 1 i 3(a) są nieparametryczne.

Definicja. *Hipotezą prostą* nazywamy hipotezę, która jednoznacznie określa rozkład badanej cechy. *Hipotezą złożoną* nazywamy hipotezę, która określa całą grupę rozkładów.

Hipoteza 2 z przykładu 3.1 jest hipotezą prostą. Pozostałe hipotezy w tym przykładzie są złożone.

W praktyce rozważamy dwie hipotezy: *hipotezę zerową* (będziemy ją oznaczać H_0) i *hipotezę alternatywną* (tą będziemy oznaczać H_1). Jeśli odrzucamy hipotezę zerową, to przyjmujemy hipotezę alternatywną i na odwrót.

Przykład 3.2.
Wiemy, że wysokość plonów przy nawożeniu starą metodą ma rozkład normalny o wartości średniej 5 *i wariancji* 1*:* $\mathcal{N}(5, 1)$ *i że wysokość plonów przy nawożeniu nową metodą ma rozkład normalny o tej samej wariancji $\sigma^2 = 1$ ale o nieznanej wartości średniej μ:* $\mathcal{N}(\mu, 1)$*. Chcemy sprawdzić czy nowa metoda zwiększyła średnią wysokość plonów.*

W tym celu będziemy testować $H_0 : \mu = 5$ *przeciwko* $H_1 : \mu > 5$.

Przeprowadzamy eksperyment losowy. Wynik takiego eksperymentu to próba czyli wektor losowy $\mathbb{X} = (X_1, X_2, \ldots, X_n)$ o wartościach w $\mathbb{R}^n$.

Definicja. *Statystyką testową* nazywamy funkcję próby $\delta(X_1, X_2, \ldots, X_n)$, która służy do weryfikacji H_0 przeciwko H_1. Zbiór wszystkich możliwych wartości funkcji δ dzielimy na dwa rozłączne zbiory W i W' takie, że:

- jeśli $\delta(x_1, x_2, \ldots, x_n) \in W$, to H_0 odrzucamy,
- jeśli $\delta(x_1, x_2, \ldots, x_n) \in W'$, to H_0 przyjmujemy.

W nazywamy *zbiorem krytycznym testu* (zbiorem odrzuceń H_0).

Musimy dobrze skonstruować statystykę testową i rozsądnie dobrać zbiór krytyczny, tak by podejmować decyzje zgodne z rzeczywistością. Jednak, ponieważ decyzje podejmujemy jedynie na podstawie próby, nie mamy całkowitej informacji o badanej populacji i w związku z tym zawsze jesteśmy narażeni na popełnienie błędu - podjęcie decyzji niezgodnej z rzeczywistością. Dokładniej, możemy popełnić jeden z dwóch błędów:

1. odrzucić H_0 w sytuacji, gdy jest ona prawdziwa (tzw. *błąd pierwszego rodzaju*);
2. przyjąć H_0 w sytuacji, gdy jest ona fałszywa (tzw. *błąd drugiego rodzaju*).

Chcielibyśmy by prawdopodobieństwa obu tych błędów były jak najmniejsze. Niestety, gdy przy ustalonej statystyce testowej, zmieniamy W tak by malał błąd pierwszego rodzaju, to błąd drugiego rodzaju rośnie i na odwrót. Postępujemy zatem tak:

- ustalamy z góry maksymalną wartość prawdopodobieństwa błędu pierwszego rodzaju, oznaczamy tą wartość α i nazywamy ją *poziomem istotności testu* (zwyczajowo przyjmuje się $\alpha = 0,01$ lub $\alpha = 0,05$, czasami $\alpha = 0,1$);
- zbiór krytyczny W wyznaczamy tak by prawdopodobieństwo błędu pierwszego rodzaju nie przekraczało α i by prawdopodobieństwo błędu drugiego rodzaju było możliwie najmniejsze.

Zapisujemy to symbolicznie:

$$\underbrace{P(\delta(X_1, X_2, \ldots, X_n) \in W | H_0)}_{\text{prawdop. błędu pierwszego rodzaju}} \quad \leq \quad \alpha \qquad \text{i}$$

$$\underbrace{P(\delta(X_1, X_2, \ldots, X_n) \notin W | H_1)}_{\text{prawdop. błędu drugiego rodzaju}} \quad - \quad \text{możliwie najmniejsze}$$

Definicja. *Moc testu* parametrycznego (*power of the test*) to funkcja zmiennej θ (gdzie θ to badany parametr) dana wzorem

$$\begin{aligned} \beta(\theta) &= P(\underbrace{\delta(X_1, X_2, \ldots, X_n)}_{\text{statystyka testowa}} \in W | \theta) = \\ &= \text{prawdop. odrzucenia } H_0 \text{ w sytuacji, gdy nieznany parametr przyjmuje wartość } \theta. \end{aligned}$$

W szczególności, jeśli $H_0 : \theta = \theta_0$, to

$$\begin{aligned} \beta(\theta_0) &= P(\delta(X_1, X_2, \ldots, X_n) \in W | \theta_0) = P(\delta(X_1, X_2, \ldots, X_n) \in W | H_0) = \\ &= \text{prawdopodobieństwo błędu pierwszego rodzaju.} \end{aligned}$$

Natomist jeśli $H_1 : \theta = \theta_1$, to wówczas

$$\begin{aligned} \beta(\theta_1) &= P(\delta(X_1, X_2, \ldots, X_n) \in W | \theta_1) = P(\delta(X_1, X_2, \ldots, X_n) \in W | H_1) = \\ &= 1 - P(\delta(X_1, X_2, \ldots, X_n) \notin W | H_1) = \\ &= 1 - \text{ prawdopodobieństwo błędu drugiego rodzaju.} \end{aligned}$$

Definicja. Najmniejszy poziom istotności, przy którym zaobserwowana wartość statystyki testowej prowadzi do odrzucenia H_0, nazywamy *p-wartością* (p-value) przeprowadzonego testu. Tzn.

$$\begin{aligned} p - \text{value} \leq \alpha \quad &\Rightarrow \quad \text{odrzucamy } H_0, \\ p - \text{value} > \alpha \quad &\Rightarrow \quad \text{przyjmujemy } H_0. \end{aligned}$$

Uwaga! Do wyników testowania statystycznego powinniśmy podchodzić z pewną rezerwą, zwłaszcza, gdy hipoteza alternatywna H_1 jest złożona. Jeśli statystyka testowa nie wpadnie do zbioru krytycznego, to stwierdzamy, że nie ma podstaw do odrzucenia H_0, co jescze nie oznacza, że H_0 należy uznać za prawdziwą.

Podstawowe testy parametryczne dla jednej populacji

Podstawowe statystyki próbkowe i funkcje w R wyliczające te statystyki	
średnia z próby: $\overline{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$, > mean(dane)	wariancja z próby: $s^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \overline{x})^2$ > var(dane)

Oznaczenia podstawowych kwantyli i funkcje w R wyliczające te kwantyle		
dla rozkładu normalnego	dla rozkładu t-Studenta	dla rozkładu chi-kwadrat
u_α	$t_{\alpha,n}$	$\chi^2_{\alpha,n}$
> qnorm(α)	> qt(α, n)	> qchisq(α, n)

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

Weryfikacje hipotez dotyczących wartości średniej na poziomie istotności α		
Model I. $X \sim N(\mu, \sigma^2), \mu$ - nieznane, σ - znane. Hipoteza zerowa $H_0 : \mu = \mu_0$. Statystyka testowa $U = \frac{\overline{X}-\mu_0}{\sigma}\sqrt{n}$.		
Hipoteza alternatywna $H_1 : \mu \neq \mu_0$ Zbiór krytyczny $W = \left(-\infty; -u_{1-\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}; +\infty\right)$	Hipoteza alternatywna $H_1 : \mu > \mu_0$ Zbiór krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Hipoteza alternatywna $H_1 : \mu < \mu_0$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha}\rangle$
Model II (t.test). $X \sim N(\mu, \sigma^2), \mu$ - nieznane, σ -nieznane. Hipoteza zerowa $H_0 : \mu = \mu_0$. Statystyka testowa $T = \frac{\overline{X}-\mu_0}{s}\sqrt{n}$.		
Hipoteza alternatywna $H_1 : \mu \neq \mu_0$ Zbiór krytyczny $W = \left(-\infty; -t_{1-\alpha/2,n-1}\right\rangle \cup \left\langle t_{1-\alpha/2,n-1}; +\infty\right)$	Hipoteza alternatywna $H_1 : \mu > \mu_0$ Zbiór krytyczny $W = \langle t_{1-\alpha,n-1}; +\infty)$	Hipoteza alternatywna $H_1 : \mu < \mu_0$ Zbiór krytyczny $W = (-\infty; -t_{1-\alpha,n-1}\rangle$
Model III. X ma rozkład dowolny (próba duża: $n \geq 100$). Hipoteza zerowa $H_0 : \mu = \mu_0$. Statystyka testowa $U = \frac{\overline{X}-\mu_0}{s}\sqrt{n}$.		
Hipoteza alternatywna $H_1 : \mu \neq \mu_0$ Zbiór krytyczny $W = \left(-\infty; -u_{1-\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}; +\infty\right)$	Hipoteza alternatywna $H_1 : \mu > \mu_0$ Zbiór krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Hipoteza alternatywna $H_1 : \mu < \mu_0$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha}\rangle$
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 $W = \left(-\infty; -u_{1-\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}; +\infty\right)$	Hipoteza alternatywna $H_1 : p > p_0$ Zbiór krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Hipoteza alternatywna $H_1 : p < p_0$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha}\rangle$
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 α		
Model. $X \sim N(\mu, \sigma^2), \mu$ - nieznane, σ - 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 $W = \left(0, \chi^2_{\alpha/2;n-1}\right\rangle \cup \left\langle \chi^2_{1-\alpha/2;n-1}; +\infty\right)$	Hipoteza alternatywna $H_1 : \sigma^2 > \sigma_0^2$ Zbiór krytyczny $W = \left\langle \chi^2_{1-\alpha;n-1}; +\infty\right)$	Hipoteza alternatywna $H_1 : \sigma^2 < \sigma_0^2$ Zbiór krytyczny $W = \left(0; \chi^2_{\alpha;n-1}\right\rangle$

Przykład 3.1. Ogrodnik ma 5000 nasion białych i czerwonych tulipanów. Chciałby wiedzieć jaki procent owych nasion to nasiona tulipanów białych. Nasiona te przeznaczone są do sprzedaży, więc nie może ich wszystkich wysiać i sprawdzić, ile z nich zakwitnie na biało. Wybrał zatem losowo 100 nasion, posiał je i okazało się, że 13 z nich ma białe kwiaty.
(a) Czy na poziomie istotności 0,01 ogrodnik może stwierdzić, że nasiona białych tulipanów stanowią 10% wszystkich nasion?
(b) Czy zmieni się odpowiedź w punkcie (a) jeśli ogrodnik posieje jedynie 10 nasion i 2 z nich wykiełkują na biało?

Przykład 3. 2. Dział kontroli jakości w zakładach chemicznych chce oszacować średnią wagę proszku do prania sprzedawanego w pudełkach o nominalnej wadze 3 kg. Pobrano w tym celu próbkę losową 7 pudełek i otrzymano wyniki (w kg): 2.93, 2.97, 3.05, 2.91, 3.02, 2.87, 2.92. Wiadomo, że rozkład wagi pudełka do prania jest normalny.
(a) Czy na poziomie istotności 0,05 można twierdzić, że faktyczna średnia waga pudełka proszku do prania jest mniejsza niż 3 kg?
(b) Zakładając, że rzeczywista średnia waga pudełka proszku do prania wynosi 2,9 kg, wyznaczyć prawdopodobieństwo, że przeprowadzając test na poziomie istotności 0.05 i na podstawie 7 obserwacji, błędnie uznamy, że średnia waga pudełka jest zgodna z podaną na pudełku.
(c) Jak liczną próbkę trzeba by pobrać, by przeprowadzony test (na poziomie istotności 0.05), w sytuacji, gdy rzeczywista średnia waga pudełka proszku do prania wynosi 2,9 kg, odrzucał hipotezę, że średnia waga pudełka jest zgodna z podaną na pudełku, z prawdopodobieństwem nie mniejszym niż 0.9.

Przykład 3. 3. Czas montowania bębna w pralce jest zmienną losową o rozkładzie normalnym z odchyleniem standardowym równym pół minuty. Norma techniczna przewiduje na tę czynność 6 minut. Wśród załogi panuje jednak przekonanie, że ten normatywny czas jest zbyt krótki. Zmierzono czas montowania bębna przez 6 losowo wybranych robotników i otrzymano następujące wyniki (w minutach): 6.2, 7.1, 6.3, 5.9, 5.5, 7.0.
(a) Na poziomie istotności 0.05 stwierdzić, czy przekonanie załogi jest słuszne.
(b) Zakładając, że rzeczywisty czas montowania bębna to 6 i ćwierć minuty, wyznaczyć prawdopodobieństwo, że przeprowadzając test na poziomie istotności 0.05 i na podstawie 6 obserwacji, błędnie uznamy, że załoga nie ma racji.

Przykład 3.4. Otrzymano następujące wyniki pomiarów grubości 6 wylosowanych detali wyprodukowanych przez zakupiony agregat (w mm.): 1.6, 1.7, 1.4, 1.5, 1.9, 1.5. Zakładamy, że rozkład grubości tego detalu jest normalny. Na poziomie istotności 0.05 zweryfikować hipotezę, że wariancja grubości detalu wykonanego przez agregat przekracza 0.03 mm^2.

Podstawowe testy parametryczne dla dwóch populacji

Przykład 3.5. 20 spośród 100 losowo wybranych studentów studiów zaocznych i 40 spośród 120 losowo wybranych studentów studiów dziennych zdało egzamin ze Statystyki w pierwszym terminie.
(a) Czy na podstawie powyższych danych możemy stwierdzić, że studenci studiów zaocznych gorzej przygotowują się do egzaminu ze Statystyki niż studenci dzienni? Przyjąć poziom istotności 0.01.
(b) Przypuszczamy, że zdawalność egzaminu ze Statystyki w pierwszym terminie wynosi dla studentów studiów zaocznych 0.2 a dziennych - 0.3. Ilu studentów studiów zaocznych i ilu studentów studiów dziennych trzeba by wylosować do próby by jednostronny test porównujący proporcje z poziomem istotności 0.01 miał moc 0.75.

Przykład 3. 6. Losową grupę 5 osób poddano 6-tygodniowej diecie odchudzającej. Uzyskano następujące wyniki (waga przed i po kuracji [w kg]):

Przed kuracją	88	86	82	64	59
Po kuracji	75	76	83	65	58

.

Można założyć, że rozkład łączny wagi przed i po kuracji jest normalny.
(a) Czy powyższe wyniki potwierdzają skuteczność diety (przyjąć poziom istotności 0.05)?
(b) Przypuszczamy, że średnia różnica wagi sprzed i po kuracji wynosi 4 kg. Jakie jest prawdopodobieństwo, że test z punktu (a) potwierdzi skuteczność diety?
(c) Przypuszczamy, że średnia różnica wagi sprzed i po kuracji wynosi 4 kg. Ile osób trzeba by losowo wybrać do eksperymentu by test jednostronny o poziomie istotności 0.05 z prawdopodobieństwem 0.8 potwierdzał skuteczność diety?

Przykład 3.7. Dokonano po 10 pomiarów tego samego napięcia prądu przy użyciu dwóch różnych woltomierzy. Dla pierwszego woltomierza otrzymano następujące wyniki

$$1.2, 1.0, 1.1, 1.4, 1.1, 1.2, 1.0, 0.9, 1.1, 1.2,$$

a dla drugiego:

$$1.3, 1.1, 1.4, 0.9, 1.4, 1.2, 1.3, 1.0, 1.2, 1.3.$$

Można założyć, że pomiary napięcia na badanych woltomierzach mają rozkłady normalne.
(a) Na poziomie istotności 0.01 zweryfikować hipotezę o jednakowych wynikach pomiaru napięcia przez oba woltomierze.
(b) Przypuszczamy, że średnia różnica pomiarów na obu woltomierzach to 0.1. Ile pomiarów na każdym woltomierzu należy wykonać by moc dwustronnego testu o poziomie istotności 0.01 wynosiła nie mniej niż 0.8.

Przykład 3.8. Pełnomocnik rządu Alfalandii d/s równego statusu kobiet i mężczyzn podejrzewa, że mężczyźni pracujący jako modele zarabiają mniej niż kobiety modelki. Czy na poziomie istotności 0,01 można uznać to podejrzenie za słusznene, jeśli średni miesięczny dochód 100 losowo wybranych modeli wyniósł 480 dukatów z odchyleniem standardowym 60 dukatów, a średni miesięczny dochód 100 losowo wybranych modelek to 600 dukatów z odchyleniem standardowym 200 dukatów?

<table>
<tr><th colspan="2">Weryfikacja hipotezy dotyczącej równości dwóch wariancji na poziomie istotności α</th></tr>
<tr><td colspan="2">UWAGA: jeżeli wyznaczona wartość statystyki F należy do W, to H_0 odrzucamy.</td></tr>
<tr><td colspan="2">Model I. (test F: var.test)
$X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, μ_1, μ_2, σ_1, σ_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).</td></tr>
<tr><td>Hipoteza alternatywna $H_1 : \sigma_1^2 \neq \sigma_2^2$
Zbiór krytyczny
$W = \langle F(1-\alpha/2, n_1-1, n_2-2); +\infty)$</td><td>Hipoteza alternatywna $H_1 : \sigma_1^2 > \sigma_2^2$
Zbiór krytyczny
$W = \langle F(1-\alpha, n_1-1, n_2-2); +\infty)$</td></tr>
<tr><td colspan="2">Powyżej $F(\alpha, n, m)$ oznacza kwantyl rozkładu F-Snedecora: $<$ qf(α, n, m)</td></tr>
</table>

Weryfikacje hipotez dotyczących dwóch średnich na poziomie istotności α		
UWAGA: jeżeli wyznaczone wartości statystyk (U lub T) należą do odpowiednich zbiorów krytycznych, to H_0 **odrzucamy**.		
Model I. $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, μ_1, μ_2 - nieznane, σ_1, σ_2 - znane; dysponujemy niezależnymi próbami losowymi z tych populacji. Hipoteza zerowa $H_0 : \mu_1 = \mu_2$. Statystyka testowa $U = \frac{\overline{x}-\overline{y}}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$.		
Hipoteza alternatywna $H_1 : \mu_1 \neq \mu_2$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha/2}\rangle \cup \langle u_{1-\alpha/2}; +\infty)$	Hipoteza alternatywna $H_1 : \mu_1 > \mu_2$ Zbiór krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Hipoteza alternatywna $H_1 : \mu_1 < \mu_2$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha}\rangle$
Model II.(unpaired t-test: t.test(...,paired=FALSE,var.equal=TRUE)) $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, μ_1, μ_2 - nieznane, σ_1, σ_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 $T = \frac{\overline{x}-\overline{y}}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$.		
Hipoteza alternatywna $H_1 : \mu_1 \neq \mu_2$ Zbiór krytyczny $W = (-\infty, -t_{1-\alpha/2, n_1+n_2-2}\rangle \cup \langle t_{1-\alpha/2, n_1+n_2-2}; +\infty)$	Hipoteza alternatywna $H_1 : \mu_1 > \mu_2$ Zbiór krytyczny $W = \langle t_{1-\alpha, n_1+n_2-2}; +\infty)$	Hipoteza alternatywna $H_1 : \mu_1 < \mu_2$ Zbiór krytyczny $W = (-\infty, -t_{1-\alpha, n_1+n_2-2}\rangle$
Jeśli 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=FALSE,var.equal=FALSE)**		
Model III.(paired t-test: t.test(..., paired=TRUE)) $(X, Y) \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, μ_1, μ_2 - nieznane, σ_1, σ_2 lub ρ - nieznane. Dysponujemy parami obserwacji, gdzie pary są wzajemnie niezależne. Hipoteza zerowa $H_0 : \mu_1 = \mu_2$. Statystyka testowa $T = \frac{\overline{z}}{s_z}\sqrt{n}$, gdzie $z_i = x_i - y_i, i = 1, 2, \ldots, n$.		
Hipoteza alternatywna $H_1 : \mu_1 \neq \mu_2$ Zbiór krytyczny $W = (-\infty; -t_{1-\alpha/2, n-1}\rangle \cup \langle t_{1-\alpha/2, n-1}; +\infty)$	Hipoteza alternatywna $H_1 : \mu_1 > \mu_2$ Zbiór krytyczny $W = \langle t_{1-\alpha, n-1}; +\infty)$	Hipoteza alternatywna $H_1 : \mu_1 < \mu_2$ Zbiór krytyczny $W = (-\infty; -t_{1-\alpha, n-1}\rangle$
Model IV. Cechy X, Y mają rozkłady dowolne ($n_1 \geq 100$, $n_2 \geq 100$), μ_1, μ_2, σ_1, σ_2 - nieznane; dysponujemy niezależnymi próbami losowymi z tych populacji. Hipoteza zerowa $H_0 : \mu_1 = \mu_2$. Statystyka testowa $U = \frac{\overline{x}-\overline{y}}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$.		
Hipoteza alternatywna $H_1 : \mu_1 \neq \mu_2$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha/2}\rangle \cup \langle u_{1-\alpha/2}; +\infty)$	Hipoteza alternatywna $H_1 : \mu_1 > \mu_2$ Zbiór krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Hipoteza alternatywna $H_1 : \mu_1 < \mu_2$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha}\rangle$
Model V (prop.test). 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 $n_1(1 - \hat{p}_1) \geq 5$ i $n_2\hat{p}_2 \geq 5$ i $n_2(1 - \hat{p}_2) \geq 5$. Hipoteza zerowa $H_0 : p_1 = p_2$. Statystyka testowa $U = \frac{\hat{p}_1-\hat{p}_2}{\sqrt{\frac{\overline{p}\overline{q}}{n}}}$, gdzie $\hat{p}_1 = \frac{k_1}{n_1}$, $\hat{p}_2 = \frac{k_2}{n_2}$, $\overline{p} = \frac{k_1+k_2}{n_1+n_2}$, $\overline{q} = 1- \overline{p}$, $n = \frac{n_1 n_2}{n_1+n_2}$.		
Hipoteza alternatywna $H_1 : p_1 \neq p_2$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha/2}\rangle \cup \langle u_{1-\alpha/2}; +\infty)$	Hipoteza alternatywna $H_1 : p_1 > p_2$ Zbiór krytyczny $W = \langle u_{1-\alpha}; +\infty)$	Hipoteza alternatywna $H_1 : p_1 < p_2$ Zbiór krytyczny $W = (-\infty; -u_{1-\alpha}\rangle$
Jeśli 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.		