

Rzeczpospolita
PolskaPolitechnika
WarszawskaUnia Europejska
Europejski Fundusz Społeczny

Rachunek Prawdopodobieństwa i Elementy Statystyki Matematycznej

Anna Dembińska

Wydział Matematyki i Nauk Informacyjnych

Wykład 11

Projekt „NERW 2 PW. Nauka – Edukacja – Rozwój – Współpraca”
współfinansowany jest ze środków Unii Europejskiej w ramach
Europejskiego Funduszu Społecznego.

Zadanie 10 pn. „Modyfikacja programów studiów na kierunkach
prowadzonych przez Wydział Matematyki i Nauk Informacyjnych”,
realizowane w ramach projektu „NERW 2 PW. Nauka – Edukacja –
Rozwój – Współpraca”, współfinansowanego ze środków Unii
Europejskiej w ramach Europejskiego Funduszu Społecznego.

11. WERYFIKACJA HIPOTEZ DOTYCZĄCYCH JEDNEJ POPULACJI

Weryfikacje hipotez dotyczących wartości średniej na poziomie istotności α		
Model I. $X \sim N(\mu, \sigma^2)$, μ - nieznane, σ - 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). $X \sim N(\mu, \sigma^2)$, μ - nieznane, σ - 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 α		
Model. $X \sim N(\mu, \sigma^2)$, μ - 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 $W = (0, \chi_{\alpha/2, n-1}^2) \cup (\chi_{1-\alpha/2, n-1}^2; +\infty)$	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 $W = (0; \chi_{\alpha, n-1}^2)$

WAŻNA UWAGA: W momencie, gdy stwierdzimy, że do rozważanego problemu pasuje nam model III z tabeli *Weryfikacje hipotez dotyczących wartości średniej* to, tak samo jak dla modelu II, możemy używać funkcji `t.test()` i `power.t.test()`. Wynika to stąd, że dla dużych n mamy $t_{\alpha,n} \approx u_{\alpha}$.

Przykład 11.1. 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. Na poziomie istotności 0.05 stwierdzić, czy przekonanie załogi jest słuszne.

Rozwiązanie przykładu 11.1:

Oznaczmy: X - czas montowania bębna w pralce. Z treści zadania wiemy, że X ma rozkład normalny ze znanym odchyleniem standardowym: $\sigma = 0,5$ min. Zapisujemy to następująco

$$X \sim \mathcal{N}(\mu, \sigma^2), \text{ gdzie } \mu \text{ nieznane, zaś } \sigma = 0,5.$$

Interesuje nas weryfikacja

$$H_0 : \mu = 6 \text{ min}$$

przeciwko

$$H_1 : \mu > 6 \text{ min.}$$

Widzimy, że do analizowanego problemu pasuje model oznaczony jako model I w tabeli Weryfikacje hipotez dotyczących wartości średniej. Aby przeprowadzić opisany tam test, wyznaczmy wartość statystyki testowej i zbiór krytyczny a następnie sprawdzimy czy statystyka testowa należy do zbioru krytycznego. Zaczniemy od statystyki testowej:

$$U = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}.$$

Obliczenia przeprowadzimy w R:

```
> czas <- c( 6.2, 7.1, 6.3, 5.9, 5.5, 7.0)
```

```
> (mean(czas)-6)/0.5*sqrt(6)
```

Otrzymujemy $u \approx 1,633$. Teraz przejdźmy do wyznaczenia zbioru krytycznego:

$$W = \langle u_{1-\alpha}; +\infty \rangle.$$

Mamy poziom istotności $\alpha = 0,05$, stąd $1 - \alpha = 0,95$ i $u_{1-\alpha} = u_{0,95}$:

```
> qnorm(0.95)
```

Otrzymujemy $u_{1-\alpha} \approx 1,645$, co daje $W \approx \langle 1,645; +\infty \rangle$. Pozostaje wyciągnąć wniosek.

$$u \approx 1,633 \notin W \approx \langle 1,645; +\infty \rangle \Rightarrow \text{brak podstaw do odrzucenia } H_0.$$

Przekonanie załogi nie jest słuszne.

Przykład 11.2. 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 białe. 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 białe?

Rozwiązanie przykładu 11.2:

Dane nasiono może być nasieniem tulipana białego lub czerwonego. Mamy zatem do czynienia z rozkładem dwupunktowym. Oznaczmy:

$$X = \begin{cases} 1 & \text{jeśli wybrane nasiono to nasiono białego tulipana,} \\ 0 & \text{jeśli wybrane nasiono to nasiono czerwonego tulipana,} \end{cases}$$

zaś p niech będzie prawdopodobieństwem trafienia na nasiono tulipana białego, tzn. $p = P(X = 1)$.

(a) $\hat{p} = \frac{k}{n} = \frac{\text{liczba sukcesów}}{\text{liczba prób}} = \frac{13}{100} = 0,13$. Sukcesem jest wylosowanie nasiona tulipana białego, bo oznaczyliśmy że $X = 1$ właśnie wtedy, gdy wybrane nasiono to nasiono białego tulipana.

Teraz szukamy w tabeli zatytuowanej Weryfikacje hipotez dotyczących wartości średniej, modelu, który pasuje do naszej sytuacji. Jest to model oznaczony numerem IV. Stawiamy hipotezy:

$$H_0 : p = 0,1$$

$$H_1 : p > 0,1 \text{ (wybrałam wersję } >, \text{ bo } \hat{p} = 0,13 \text{ sugeruje, że } p \text{ może być większe niż } 0,1.)$$

Zauważmy, że

- $n\hat{p} = 13 \geq 5$ ($n\hat{p}$ to liczba sukcesów, czyli liczba wylosowanych nasion tulipanów białych),
- $n\hat{q} = 87 \geq 5$ ($\hat{q} = 1 - \hat{p}$, więc $n\hat{q}$ to liczba porażek, czyli liczba wylosowanych nasion tulipanów czerwonych).

Zatem możemy użyć `prop.test()`, czyli testu, w którym rozkład statystyki testowej jest przybliżany rozkładem normalnym:

```
> prop.test(x=13,n=100,p=0.1,alternative="greater")
```

Powyżej argument `x` oznacza liczbę otrzymanych sukcesów, a `n` - liczbę wszystkich prób. Odczytujemy p wartość:

$$p\text{-value} = 0,2023 > \alpha = 0,01 \Rightarrow \text{nie ma podstaw do odrzucenia } H_0,$$

gdzie $\alpha = 0,01$ to poziom istotności testu. Wyciągamy więc wniosek, że ogrodnik może stwierdzić, że nasiona białych tulipanów stanowią 10% wszystkich nasion.

(b) Nadal testujemy $H_0 : p = 0,1$ przeciwko $H_1 : p > 0,1$. Jednak teraz $n = 10$ i $k = 2$, więc mamy

$$n\hat{p} = 2 \not\geq 5$$

i nie możemy zastosować `prop.test()` (wynika to stąd, że nie będzie działać przybliżenie rozkładem normalnym; aby ono działało muszą jednocześnie być spełnione oba warunki: $n\hat{p} \geq 5$ i $n\hat{q} \geq 5$). W tej sytuacji należy użyć testu dokładnego `binom.test()`:

```
> binom.test(x=2,n=10,p=0.1,alternative="greater")
```

$p\text{-value} = 0,2639 > \alpha = 0,01 \Rightarrow$ nie ma podstaw do odrzucenia H_0 ,

zatem odpowiedź z punktu (a) nie ulega zmianie: ogrodnik może stwierdzić, że nasiona białych tulipanów stanowią 10% wszystkich nasion.

Przykład 11.3. 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.

Rozwiązanie przykładu 11.3:

Oznaczmy: X -waga proszku do prania. Z treści zadania wiemy, że X ma rozkład normalny, ale parametrów tego rozkładu nie znamy. Zapisujemy to natępująco

$$X \sim \mathcal{N}(\mu, \sigma^2), \text{ gdzie } \mu \text{ i } \sigma \text{ są nieznane.}$$

Stąd widzimy, że będzie nam pasować model oznaczony jako model II w tabeli Weryfikacje hipotez dotyczących wartości średniej.

Wpisujemy dane do R:

```
> waga.proszku <- c(2.93, 2.97, 3.05, 2.91, 3.02, 2.87, 2.92)
```

(a) Stawiamy hipotezy:

$$H_0 : \mu = 3 \text{ kg}$$

$$H_1 : \mu < 3 \text{ kg}$$

Do weryfikacji H_0 przeciwko H_1 użyjemy `t.test`:

```
> t.test(x=waga.proszku, alternative="less", mu=3)
```

$$p\text{-value} = 0,04952 < \alpha = 0,05 \Rightarrow \text{odrzucaamy } H_0.$$

Zatem uznajemy, że rzeczywiście faktyczna średnia waga pudełka proszku do prania jest mniejsza niż 3 kg.

Powyższy test można przeprowadzić także w inny sposób: wyznaczając wartość statystyki testowej i zbiór krytyczny oraz sprawdzając czy statystyka testowa należy do zbioru krytycznego. Zaczniemy od statystyki testowej:

$$T = \frac{\bar{X} - \mu_0}{s} \sqrt{n},$$

gdzie \bar{X} to średnia z próby, s to odchylenie standardowe z próby, zaś n to liczność próby. Do rachunków użyjemy R:

```
> (mean(waga.proszku)-3)/sd(waga.proszku)*sqrt(7)
```

Otrzymujemy $t \approx -1,95$. Przechodzimy do wyznaczenia zbioru krytycznego W :

$$W = (-\infty; -t_{1-\alpha, n-1}),$$

gdzie $\alpha = 0,05$ to poziom istotności, co daje $1 - \alpha = 0,95$. Kwantyl $t_{1-\alpha, n-1}$ wyznaczymy przy pomocy R:

```
> qt(0.95, 7-1)
```

Mamy $t_{1-\alpha, n-1} \approx 1,943$, co daje $W \approx (-\infty; -1,943)$. Widzimy, że

$$t \approx -1,95 \in W \approx (-\infty; -1,943),$$

więc odrzucaamy H_0 i stwierdzamy, że faktyczna średnia waga pudełka proszku do prania jest mniejsza niż 3 kg.

(b) Zakładamy, że $\mu = \mu_1 = 2,9$. Przy tym założeniu chcemy policzyć prawdopodobieństwo, że uznamy, że średnia waga pudełka jest zgodna z podaną na pudełku. Zatem szukamy prawdopodobieństwa, że przyjmiemy H_0 w sytuacji, gdy wartość badanego parametru to 2,9:

$$P(\text{przyjmiemy } H_0 | \mu = 2,9) = ?$$

Przypomnijmy, że

$$\text{moc.testu}(\beta) = P(\text{odrzućmy } H_0 | \text{badany parametr} = \beta).$$

Stąd

$$P(\text{przyjmiemy } H_0 | \mu = 2,9) = 1 - P(\text{odrzućmy } H_0 | \mu = 2,9) = 1 - \text{moc.testu}(2,9).$$

Użyjemy funkcji `power.t.test()`, która jest związana z mocą t -testu. Ma ona następujące argumenty:

- `power` - moc testu,
- `n` - liczność próby,
- `delta` = $|\mu_0 - \mu_1|$,
- `sd` - odchylenie standardowe badanej cechy (tutaj wagi pudełka proszku do prania), nie znamy go, więc go przybliżamy odchyleniem standardowym z próby, mając jednak świadomość, że doprowadzi to nas do wyniku przybliżonego; argument `sd` jest domyślnie ustawiony na 1,
- `sig.level` - poziom istotności testu, domyślnie ustawiony na 0.05,
- `type` - mamy do wyboru `type="one.sample"`, `"two.sample"` lub `"paired"`; na razie zajmujemy się testami dla jednej populacji, więc wybieramy `"one.sample"`,
- `alternative` - mamy do wyboru `alternative="one.sided"` lub `"two.sided"`:
 - `"one.sided"` używamy, gdy H_1 jest postaci $H_1 : \mu < \mu_0$ lub $H_1 : \mu > \mu_0$,
 - `"two.sided"` używamy, gdy H_1 jest postaci $H_1 : \mu \neq \mu_0$.

Jeden z pięciu pierwszych wyżej wymienionych argumentów funkcji `power.t.test()` musimy zostawić pusty i właśnie ten argument zostanie wyliczony. Chcąc wyznaczyć `sd` lub `sig.level` należy napisać `sd=NULL` lub odpowiednio `sig.level=NULL`, aby do tych argumentów nie została automatycznie przypisana ich wartość domyślna.

W celu rozwiązania naszego problemu napiszemy (`delta = |\mu_0 - \mu_1| = |3 - 2,9| = 0,1`):

```
> power.t.test(n=7, delta=0.1, sd=sd(waga.proszku), sig.level=0.05,
               type="one.sample", alternative="one.sided")
```

R wypisze podane przez nas wartości argumentów i wyliczoną wartość mocy testu (`power`). Aby uzyskać szukane prawdopodobieństwo musimy od 1 odjąć wyliczoną wartość mocy testu. Możemy to zrobić automatycznie pisząc

```
> 1-power.t.test(n=7, delta=0.1, sd=sd(waga.proszku), sig.level=0.05,
                 type="one.sample", alternative="one.sided")$power
```

Otrzymujemy 0,0241. Jest to prawdopodobieństwo przyjęcia $H_0 : \mu = 3$ kg w sytuacji, gdy $\mu = 2,9$ kg czyli prawdopodobieństwo popełnienia błędu. Zatem dobrze, że jest całkiem małe.

(c) Nadal zakładamy, że $\mu = \mu_1 = 2,9$. Przy tym założeniu szukamy n takiego by prawdopodobieństwo odrzucenia H_0 było nie mniejsze niż 0,9:

szukamy n takiego by $P(\text{odrzućmy } H_0 | \mu = 2,9) \geq 0,9$

czyli

szukamy n takiego by $\text{moc.testu}(2,9) \geq 0,9$.

Użyjemy funkcji `power.t.test()`:

```
> power.t.test(power=0.9, delta=0.1, sd=sd(waga.proszku),
               sig.level=0.05, type="one.sample",
               alternative="one.sided")
```

Otrzymujemy $n = 5,186$ co oznacza, że potrzebujemy próbkę o liczności $n = 6$ (zaokrąglamy do góry, aby moc testu nie spadła poniżej 0,9).

Przykład 11.4. 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 .

Rozwiązanie przykładu 11.4:

Oznaczmy: X - grubość detalu. Z treści zadania wiemy, że X ma rozkład normalny, ale parametrów tego rozkładu nie znamy. Zapisujemy to natępująco

$$X \sim \mathcal{N}(\mu, \sigma^2), \text{ gdzie } \mu \text{ i } \sigma \text{ są nieznane.}$$

Interesuje nas weryfikacja

$$H_0 : \sigma^2 = 0,03 \text{ mm}^2$$

przeciwko

$$H_1 : \sigma^2 > 0,03 \text{ mm}^2.$$

Powyższe hipotezy dotyczą wariancji. Patrzymy więc na dolną tabelę Weryfikacja hipotezy dotyczącej jednej wariancji. Przedstawiony w niej jeden model pasuje do naszej sytuacji. Aby przeprowadzić opisany tam test, wyznaczmy wartość statystyki testowej i zbiór krytyczny a następnie sprawdzimy czy statystyka testowa należy do zbioru krytycznego. Zaczniemy od statystyki testowej:

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

gdzie s^2 to wariancja z próby. Rachunki wykonujemy w R:

```
> grubosc <- c(1.6, 1.7, 1.4, 1.5, 1.9, 1.5)
```

```
> (6-1)*var(grubosc)/0.03
```

Statystyka testowa $\chi^2 \approx 5,333$. Teraz zajmijmy się zbiorem krytycznym

$$W = \langle \chi_{1-\alpha; n-1}^2; +\infty \rangle.$$

Poziom istotności $\alpha = 0,05$, zatem $1 - \alpha = 0,95$. Liczymy $\chi_{1-\alpha; n-1}^2 = \chi_{0,95; 5}^2$:

```
> qchisq(0.95, 5)
```

Otrzymujemy $\chi_{1-\alpha; n-1}^2 \approx 11,0705$, co daje $W \approx \langle 11,0705; \infty \rangle$. Widzimy, że

$$\chi^2 \approx 5,333 \notin W \approx \langle 11,0705; \infty \rangle,$$

wiec nie ma podstaw do odrzucenia hipotezy zerowej. Uznajemy, że wariancja grubości detalu wykonanego przez agregat nie przekracza 0.03 mm^2 .