Hipotezy proste

Zadanie 1. Niech X ma funkcję gęstości

$$f_a(x) = \begin{cases} (1+a)x^a, & \text{dla } 0 < x < 1, \\ 0, & \text{poza tym.} \end{cases}$$

Testujemy $H_0 : a = 1$ przeciwko $H_1 : a = 2$. Dysponujemy pojedynczą obserwacją X. Wyznaczyć obszar krytyczny najmocniejszego testu o rozmiarze 0.1.
Odp. $(\sqrt{0.9}, +\infty)$

Zadanie 2. Zmienna losowa X ma gęstość prawdopodobieństwa $f(x)$. Na podstawie pojedynczej obserwacji X przeprowadzamy test hipotezy

$$H_0 : \ f(x) = \begin{cases} 1, & \text{dla } 0 < x < 1, \\ 0, & \text{poza tym,} \end{cases}$$

przeciwko alternatywie:

$$H_1 : \ f(x) = \begin{cases} 5x^4, & \text{dla } 0 < x < 1, \\ 0, & \text{poza tym.} \end{cases}$$

Wyznaczyć moc najmocniejszego testu na poziomie istotności α.
Odp. $1 - (1-\alpha)^5$

Zadanie 3. $X_1, \ldots, X_{20}$ jest próbą losową z rozkładu normalnego o parametrach $(0, \sigma^2)$. Rozważmy najmocniejszy test hipotezy $H_0 : \sigma^2 = 1$ przeciwko alternatywie $H_1 : \sigma^2 = 3$ na poziomie istotności 0.1. Wyznaczyć moc tego testu.
Odp. ≈ 0.90

Zadanie 4. Niech $X_1, \ldots, X_n$ będzie próbą prostą z rozkładu $N(\mu, 4)$. Rozważmy najmocniejszy test hipotezy $H_0 : \mu = 0$ przeciw alternatywie $H_1 : \mu = 1$, na poziomie istotności 0.01. Ile potrzeba obserwacji, żeby moc testu była większa niż 0.9?
Odp. $n \geq 53$

Zadanie 5. Wiadomo, że $X_1, \ldots, X_n$ jest próbą z rozkładu normalnego $N(\mu - \theta, 1)$, zaś $Y_1, \ldots, Y_n$ jest niezależną próbą z rozkładu $N(\mu + \theta, 1)$. Liczby μ i θ są nieznanymi parametrami. Rozpatrujemy zadanie testowania hipotezy $H_0 : \theta = 0$ przeciw alternatywie $H_1 : \theta = 1$. Dla jakich n można skonstruować test na poziomie istotności 0.05 o mocy przynajmniej 0.95?
Odp. $n \geq 22$

Zadanie 6. Swego czasu zaobserwowano realizację $x_1, \ldots, x_{20}$ prostej próby z rozkładu normalnego o nieznanej wartości oczekiwanej μ i wariancji. W celu testowania hipotezy $\mu = 0$ chcielibyśmy wykorzystać te dane. Niestety część z danych zaginęła i dziś dysponujemy jedynie obserwacjami $x_1, \ldots, x_{10}$ oraz średnią $\bar{x}_{20} = \frac{1}{20}\sum_{i=1}^{20} x_i$. Pięciu statystyków zaproponowało pięć różnych rozwiązań, każde z nich uwzględniające w ten lub ów sposób dodatkową informację o średniej $\bar{x}_{20}$. Każdy ze statystyków twierdzi, że podana przez niego statystyka testowa jest realizacją (przy założeniu $\mu = 0$) zmiennej losowej o rozkładzie Studenta z podaną liczbą stopni swobody:

Statystyka (A) $\sqrt{10}\dfrac{\bar{x}_{10}}{s_{10}}$ pochodzi z rozkładu t_9

Statystyka (B) $\sqrt{20}\dfrac{\bar{x}_{20}}{s_{20}}$ pochodzi z rozkładu t_{19}

Statystyka (C) $\sqrt{20}\dfrac{\bar{x}_{20}}{s_{10}}$ pochodzi z rozkładu t_9

Statystyka (D) $\sqrt{20}\dfrac{\bar{x}_{20}}{s_{20}}$ pochodzi z rozkładu t_9

Statystyka (E) $\sqrt{10}\dfrac{\bar{x}_{20}}{s_{20}}$ pochodzi z rozkładu t_9

$$\bar{x}_{10} = \frac{1}{10}\sum_{i=1}^{10} x_i, \quad s_{10} = \sqrt{\frac{1}{9}\sum_{i=1}^{10}(x_i - \bar{x}_{10})^2}, \quad s_{20} = \sqrt{\frac{1}{9}\sum_{i=1}^{10}(x_i - \bar{x}_{20})^2}.$$

Który ze statystyków ma rację?
Odp. C

Zadanie 7. Wyniki oszacowania trzech alternatywnych modeli rozkładu ilości szkód N na tej samej (licznej) próbce ryzyk dały następujące wartości logarytmu funkcji wiarygodności:
-571.22 dla zwykłego rozkładu Poissona (model 1);
-572.20 dla modelu ze swobodnym parametrem $P(N = 0)$ oraz ogonem Poissonowskim (model 2);
-573.25 dla modelu ze swobodnymi parametrami $P(N = 0)$ oraz $P(N = 1)$ oraz ogonem Poissonowskim (model 3).
Dokonujemy doboru modelu na podstawie testów ilorazu wiarygodności, przeprowadzając go w przypadku porównania każdej pary modeli na poziomie istotności 0.05. Który z modeli należy wybrać na tej podstawie?
Odp. podane informacje są sprzeczne!

Zadanie 8. $U_1, \ldots, U_n, \ldots$ są niezależnymi zmiennymi losowymi o rozkładzie jednostajnym na przedziale $(0, 1)$. Obserwujemy zmienną losową X i rozpatrujemy $H_0 : X$ ma rozkład taki jak $\min\{U_1, U_2, U_3\}$ przeciwko $H_1 : X$ ma rozkład taki jak $\min\{U_1, U_2\}$. Wyznaczyć moc najmocniejszego testu na poziomie istotności $1/8$.
Odp. $1/4$

Zadanie 9. Rozważamy zadanie testowania, na podstawie pojedynczej obserwacji X, hipotezy prostej $H_0 : X$*pochodzi z rozkladu* f_0 przeciwko prostej alternatywie $H_1 : X$*pochodzi z rozkladu* f_1. Wiadomo, że dla każdego $\alpha \in (0, 1)$ najmocniejszy test na poziomie istotności α odrzucający H_0, jeśli $X > k = k(\alpha)$, ma moc $1 - \beta$ taką, że $1 - \beta = 1 - \beta(\alpha) = \alpha^2$. Gęstość f_0 dla $x > 0$ określona jest wzorem $\frac{1}{(1+x)^2}$. Jakim wzorem (dla dodatnich x) określona jest gęstość f_1?
Odp. $\frac{2}{(1+x)^3}$

Zadanie 10. Urna zawiera r kul ponumerowanych liczbami $1, 2, \ldots, r$. Liczba kul r jest nieznanym parametrem, o którym wiemy, że jest większy od 5. Wybieramy z urny pięć kul, losując je bez zwracania. Na podstawie numerów wylosowanych kul testujemy hipotezę zerową $H_0 : r = 25$ przeciwko alternatywie $H_1 : r = 48$. Obliczyć moc najmocniejszego testu na poziomie istotności 0.2 (z dokładnością do trzech cyfr po kropce dziesiętnej).
Odp. 0.975

Zadanie 11. Załóżmy, że dysponujemy pojedynczą obserwacją X z rozkładu normalnego $N(\mu, \sigma^2)$. Rozważmy zadanie testowania hipotezy $H_0 : \mu = 0, \sigma^2 = 1$ przeciw alternatywie $H_1 : \mu = 1, \sigma^2 = 4$. Najmocniejszy test na poziomie istotności α jest postaci

$$\text{odrzucić } H_0, \text{ gdy } X \notin (-2, b).$$

Podać b i poziom istotności α.
Odp. $b = 4/3$, $\alpha = 0.11$

Zadanie 12. Załóżmy, że $X_1, \ldots, X_n$ jest próbką z rozkładu normalnego $N(\mu, 1)$ z nieznanym parametrem μ i znaną wariancją $\sigma^2 = 1$. Znaleźć najmniejsze n, dla którego istnieje test hipotezy $H_0 : \mu = 10.0$ przeciwko alternatywie $H_1 : \mu = 10.1$ na poziomie istotności 0.05 o mocy przynajmniej 0.50.
Odp. 271

Zadanie 13. Niech X będzie pojedynczą obserwacją z rozkładu Pareto o dystrybuancie

$$F_\lambda(x) = \begin{cases} 1 - \frac{\lambda}{\lambda + x}, & \text{dla } x > 0, \\ 0, & \text{poza tym.} \end{cases}$$

Rozważmy najmocniejszy test hipotezy $H_0 : \lambda = 1$ przeciwko alternatywie $H_1 : \lambda = 101$ na poziomie istotności 0.01. Wyznaczyć moc tego testu.
Odp. 0.505

Zadanie 14. Obserwujemy parę (X, Y) zmiennych losowych. Zakładamy, że są to zmienne niezależne, X ma rozkład normalny $N(\mu_X, 1)$ i Y ma rozkład normalny $N(\mu_Y, 1/3)$. Rozważmy najmocniejszy test hipotezy $H_0 : (\mu_X, \mu_Y) = (0, 0)$ przeciwko alternatywie $H_1 : (\mu_X, \mu_Y) = (1, 1)$ na poziomie istotności 0.1. Wyznaczyć moc tego testu.
Odp. 0.76

Zadanie 15. Niech $X_1, \ldots, X_9$ będzie próbką z rozkładu normalnego $N(\mu, 1)$ o nieznanej wartości oczekiwanej μ i znanej wariancji $\sigma^2 = 1$. Rozpatrzmy zadanie testowania hipotezy $H_0 : \mu = 0$ przeciwko alternatywie $H_1 : \mu = 0.5$. Należy zbudować taki test, dla którego suma prawdopodobieństw błędów I i II rodzaju, oznaczanych odpowiednio przez α i β jest najmniejsza. Obliczyć tę najmniejszą wartość.
Odp. 0.4533

Zadanie 16. Rozważmy następujące zagadnienie testowania hipotez statystycznych. Dysponujemy próbką $X_1, \ldots, X_n$ z rozkładu normalnego o nieznanej średniej μ i znanej wariancji równej 1. Przeprowadzamy najmocniejszy test hipotezy $H_0 : \mu = 0$ przeciwko alternatywie $H_1 : \mu = 1$ na poziomie istotności 0.5. Oczywiście, moc tego testu zależy od rozmiaru próbki. Niech β_n oznacza prawdopodobieństwo błędu drugiego rodzaju, dla rozmiaru próbki n. Udowodnić, że

$$\lim_{n\to\infty} \frac{\beta_n}{e^{-n/2}/\sqrt{2\pi n}} = 1.$$

Odp. —

Zadanie 17. Obserwujemy pary $(X_1, Y_1), \ldots, (X_{20}, Y_{20})$. Zmienne $X_1, \ldots, X_{20}$ są zmiennymi losowymi o rozkładzie normalnym $N(\mu_X, 1)$, a zmienne $Y_1, \ldots, Y_{20}$ o rozkładzie normalnym $N(\mu_Y, 4)$. Wszystkie zmienne są niezależne. Rozważamy test najmocniejszy hipotezy $H_0 : (\mu_X, \mu_Y) = (0, 0)$ przeciw alternatywie $H_1 : (\mu_X, \mu_Y) = (1, 1)$ na poziomie istotności 0.01. Wyznaczyć prawdopodobieństwo błędu drugiego rodzaju tego testu.
Odp. 0.004

Zadanie 18. Załóżmy, że dysponujemy pojedynczą obserwacją X z rozkładu Laplace'a o gęstości

$$f_{\mu,\lambda}(x) = \frac{\lambda}{2} e^{-\lambda|x-\mu|}, \quad x \in R,$$

gdzie $\lambda > 0$ i $\mu \in R$ są parametrami. Rozważmy zadanie testowania hipotezy

$$H_0 : \mu = 0 \text{ i } \lambda = 1$$

przeciw alternatywie

$$H_1 : \mu = -1 \text{ i } \lambda = 0.5$$

Obszar krytyczny najmocniejszego testu na poziomie istotności α jest postaci

$$K = \{x : x \notin (a, 3)\}.$$

Wyznaczyć a i poziom istotności α.
Odp. $a = -1$, $\alpha = 0.209$

Zadanie 19. Załóżmy, że dysponujemy pojedynczą obserwacją X z rozkładu $P \in \{P_0, P_1\}$, gdzie P_0 jest rozkładem normalnym $N(0,1)$ i P_1 jest rozkładem Laplace'a o gęstości $f(x) = \frac{1}{2}e^{-|x|}$. Rozważmy zadanie testowania hipotezy $H_0 : P = P_0$ przeciw alternatywie $H_1 : P = P_1$. Podać rozmiar testu najmocniejszego, jeśli wiadomo, że obszar krytyczny testu jest sumą przedziałów rozłącznych, z których jeden jest równy $(-\infty, -1.9)$.
Odp. 0.137

Zadanie 20. Zakładamy, że zależność czynnika Y od czynnika x (nielosowego) opisuje model regresji liniowej $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, gdzie błędy ε_i są niezależne i mają rozkłady normalne o wartości oczekiwanej 0 i wariancji 1. Obserwujemy zmienne losowe $Y_1, \ldots, Y_n$ przy danych wartościach $X_1, \ldots, X_n$. Wyznaczyć obszar krytyczny testu najmocniejszego dla weryfikacji hipotezy $H_0 : \beta_0 = 0$ i $\beta_1 = 1$ przy alternatywie $H_1 : \beta_0 = 1$ i $\beta_1 = 2$ na poziomie istotności 0.05.
Odp. $\frac{\sum_{i=1}^{n}(Y_i - x_i)(1+x_i)}{\sqrt{\sum_{i=1}^{n}(1+x_i)^2}} > 1.645$

Zadanie 21. Niech $X_1, \ldots, X_{10}$ będą niezależnymi zmiennymi losowymi o tym samym rozkładzie ciągłym o ściśle rosnącej dystrybuancie F. Hipotezę H_0 : *F jest dystrybuantą rozkładu symetrycznego*, tzn. takiego że dla każdego x zachodzi: $F(-x) = 1 - F(x)$, odrzucamy, gdy spełniona jest nierówność $K > 7$ lub $K < 3$, gdzie K jest liczbą elementów w próbie losowej $X_1, \ldots, X_{10}$ o wartościach większych od zera. Wyznaczyć rozmiar testu.
Odp. 7/64

Zadanie 22. Niech $X_1, \ldots, X_9$ będą niezależnymi zmiennymi losowymi o tym samym rozkładzie jednostajnym na przedziale $[-\theta, \theta]$, gdzie $\theta > 0$ jest nieznanym parametrem. Hipotezę $H_0 : \theta = 2$ przy alternatywie $H_1 : \theta = 4$ weryfikujemy testem najmocniejszym na poziomie istotności 0.1. Wyznaczyć prawdopodobieństwo błędu drugiego rodzaju.
Odp. 0.0035

Hipotezy złożone

Zadanie 23. Niech $X_1, \dots, X_n$ będzie próbą z rozkładu gamma o gęstości

$$f_p(x) = \begin{cases} \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x}, & \text{dla } x > 0, \\ 0, & \text{poza tym,} \end{cases}$$

gdzie $\lambda > 0$ jest znane, zaś $p > 0$ jest nieznane. Wyznaczyć obszar krytyczny jednostajnie najmocniejszego testu hipotezy $H_0 : p = 2$ przeciw hipotezie alternatywnej $H_1 : p > 2$ na poziomie istotności α.
Odp. $\ln X_1 + \dots + \ln X_n > k$, gdzie k zależy od parametru λ i liczby α

Zadanie 24. Niech $X_1, \dots, X_n$ będzie próbą z rozkładu jednostajnego na odcinku (a, b), gdzie $0 \le a < b$. Rozważmy zagadnienie weryfikacji hipotezy $H_0 : a = 0$ przeciw hipotezie alternatywnej $H_1 : a > 0$. Wyznaczyć obszar krytyczny testu ilorazu wiarogodności.
Odp. $\frac{\min\{X_1,\dots,X_n\}}{\max\{X_1,\dots,X_n\}} > k$

Zadanie 25. Gęstość zmiennej losowej X ma postać $f_\theta(x) = \frac{1}{2} e^{-|x-\theta|}$ dla $x \in R$, gdzie θ jest nieznanym parametrem. Rozważmy jednostajnie najmocniejszy test hipotezy $H_0 : \theta = 0$ przeciw hipotezie alternatywnej $H_1 : \theta > 0$ na poziomie istotności α, gdzie $\alpha < 0.5$, oparty na pojedynczej obserwacji X. Dla jakiego θ funkcja mocy tego testu osiąga wartość 0.75?
Odp. $-\ln \alpha$

Zadanie 26. Zakładamy, że liczba roszczeń w ciągu roku dla pewnego portfela ryzyk jest zmienną losową X o rozkładzie Poissona. Zaobserwowano $X = 2600$. W oparciu o przybliżenie rozkładu Poissona rozkładem normalnym zbudowano test hipotezy $H_0 : EX = 2500$ przeciwko alternatywie $H_1 : EX > 2500$ o obszarze krytycznym postaci $X > c$. Czy test odrzuca H_0 na poziomie istotności 0.05?
Odp. Tak

Zadanie 27. Niech X będzie pojedynczą obserwacją z przesuniętego rozkładu wykładniczego

$$f_\theta(x) = \begin{cases} e^{-(x-\theta)}, & \text{dla } x > \theta, \\ 0, & \text{poza tym,} \end{cases}$$

gdzie $\theta \ge 0$ jest nieznanym parametrem. Rozważmy jednostajnie najmocniejszy test hipotezy $H_0 : \theta = 0$ przeciwko alternatywie $H_1 : \theta > 0$, na poziomie istotności 0.05. Wyznaczyć zbiór tych wartości θ, dla których moc testu wynosi co najmniej 0.90.
Odp. $[\ln 2 - \ln 10 + \ln 9, +\infty)$

Zadanie 28. Mamy próbę $(X_1, Y_1), \dots, (X_n, Y_n)$ z dwuwymiarowego rozkładu normalnego o nieznanych parametrach:

$$EX_i = EY_i = \mu,\ VarX_i = VarY_i = \sigma^2,\ Cov(X_i, Y_i) = \varrho\sigma^2.$$

Niech $Z_i = X_i + Y_i$ oraz $R_i = X_i - Y_i$,

$$S_Z^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Z_i - \bar{Z})^2, \quad S_R^2 = \frac{1}{n-1} \sum_{i=1}^{n} (R_i - \bar{R})^2,$$

gdzie $\bar{Z}$ oraz $\bar{R}$ to odpowiednie średnie z próbki. Niech $\varrho_0 \in (-1, 1)$ będzie ustaloną liczbą, $\varrho_0 \ne 0$. Do testowania hipotezy $H_0 : \varrho = \varrho_0$ przeciwko alternatywie $H_1 : \varrho \ne \varrho_0$ możemy użyć testu opartego na statystyce S_Z^2 / S_R^2. Wyznaczyć rozkład tej statystyki przy założeniu prawdziwości hipotezy zerowej.
Odp. $\frac{1-\varrho_0}{1+\varrho_0} \frac{S_Z^2}{S_R^2}$ ma rozkład $F(n-1, n-1)$

Zadanie 29. Niech (X_1, X_2, X_3, X_4) będzie czteroelementową próbką z rozkładu normalnego $N(\mu, \sigma^2)$ o nieznanych parametrach. Testujemy hipotezę $H_0 : \sigma^2 = \sigma_0^2$ przeciw alternatywie $H_0 : \sigma^2 > \sigma_0^2$ za pomocą statystyki $\frac{\sum_{i=1}^{4}(X_i-\bar{X})^2}{\sigma_0^2}$, gdzie $\bar{X} = \frac{1}{4}\sum_{i=1}^{4} X_i$ przyjmując poziom istotności 0.05. Zaobserwowaliśmy $(1, -1.2, 3, 0.7)$. Jaka jest minimalna możliwa wartość σ_0^2, skoro wiadomo, że test wykazał, że nie mamy podstaw do odrzucenia hipotezy H_0?
Odp. 1.13

Zadanie 30. Rozważmy prosty model regresji liniowej bez wyrazu wolnego $Y_i = \theta x_i + \varepsilon_i$ $(i = 1, \ldots, n)$, gdzie (Y_i, x_i) to obserwacje par (losowa zmienna zależna, nielosowa zmienna niezależna), θ jest nieznanym parametrem, a $(\varepsilon_1, \ldots, \varepsilon_n)$ są nawzajem niezależnymi zmiennymi losowymi o identycznym rozkładzie normalnym z zerową wartością oczekiwaną i znaną wariancją σ^2. Rozważmy jednostajnie najmocniejszy test hipotezy $H_0 : \theta = 0$ przeciw alternatywie $H_1 : \theta > 0$ na poziomie istotności 0.05. Wyznaczyć obszar krytyczny tego testu.
Odp. $\sum_{i=1}^{n} Y_i x_i > u_{0.95}\sigma\sqrt{\sum_{i=1}^{n} x_i^2}$, gdzie u_γ jest kwantylem rzędu γ standardowego rozkładu normalnego

Zadanie 31. Niech $X_1, \ldots, X_n$ będzie próbą losową z rozkładu jednostajnego na przedziale $(0, \theta)$. W celu przetestowania hipotezy $H_0 : \theta = 1$ przeciw alternatywie $H_1 : \theta > 1$ wykorzystujemy test o obszarze krytycznym $\max\{X_1, \ldots, X_n\} > c$. Niech n_0 będzie najmniejszą liczebnością próbki taką, przy której test na poziomie istotności 0.20 ma w punkcie $\theta = \sqrt[3]{2}$ moc co najmniej 0.95. Obliczyć n_0.
Odp. 13

Zadanie 32. Niech $X_1, \ldots, X_n$ będzie próbką n niezależnych realizacji z rozkładu o dystrybuancie

$$F_\theta(x) = \begin{cases} 1 - 2^{-(x-\theta)}, & \text{dla } x > \theta, \\ 0, & \text{poza tym,} \end{cases}$$

gdzie $\theta > 0$ jest nieznanym parametrem. Rozważmy jednostajnie najmocniejszy test hipotezy $H_0 : \theta = 0$ przeciw alternatywie $H_1 : \theta > 0$ na poziomie istotności 0.01. Przy jakiej liczebności próbki n, w danym punkcie $\theta_1 > 0$ funkcja mocy tego testu przybiera wartość większą lub równą 0.64?
Odp. $n \geq 6/\theta_1$

Zadanie 33. Załóżmy, że $X_1, \ldots, X_{10}$ jest próbką z rozkładu normalnego $N(\mu, \sigma^2)$ ze znaną średnią μ i nieznaną wariancją σ^2. Rozważmy test jednostajnie najmocniejszy hipotezy $H_0 : \sigma^2 \leq 1$ przeciw alternatywie $H_1 : \sigma^2 > 1$ na poziomie istotności 0.05. Dla jakich wariancji σ^2 moc testu przekracza 0.9?
Odp. $\sigma^2 \geq 3.7628$

Zadanie 34. Dana jest próbka $X_1, \ldots, X_{10}$ z rozkładu normalnego $N(\mu, \sigma^2)$ z nieznanymi parametrami μ i σ^2. Rozważamy problem testowania hipotezy $H_0 : \mu = 0$ przeciw alternatywie $H_1 : \mu \neq 0$. Używamy testu, który odrzuca H_0 jeśli $|\bar{X}/V| > c$, gdzie $V^2 = \frac{1}{10}\sum_{i=1}^{10} X_i^2$. Dobrać stałą c tak, żeby prawdopodobieństwo błędu I rodzaju tego testu było równe 0.05.
Odp. 0.6021

Zadanie 35. Zmienne losowe $X_1, \ldots, X_9$ są wzajemnie niezależne, X_i ma rozkład normalny $N(\mu\sqrt{i}, 1)$ dla $i = 1, \ldots, 9$. Rozważamy hipotezy statystyczne $H_0 : \mu = 0$ i $H_1 : \mu > 0$. Znaleźć obszar krytyczny testu jednostajnie najmocniejszego hipotezy zerowej H_0 przeciw alternatywie H_1 na poziomie istotności 0.025.
Odp. $\sum_{i=1}^{9} \sqrt{i} X_i > 1.96\sqrt{45}$

Zadanie 36. Niech $X_1, \ldots, X_{10}$ będzie próbką z rozkładu prawdopodobieństwa o gęstości

$$f_\theta(x) = \begin{cases} \theta x^{\theta-1}, & \text{dla } 0 < x < 1, \\ 0, & \text{poza tym.} \end{cases}$$

Rozważmy test jednostajnie najmocniejszy hipotezy $H_0 : \theta = 1$ przeciwko alternatywie $H_1 : \theta > 1$ na poziomie istotności 0.01. Dla jakich wartości parametru θ ten test ma moc nie mniejszą niż 0.99?
Odp. $\theta \geq 4.55$

Zadanie 37. Próbka $X_1, \dots, X_n$ pochodzi z rozkładu normalnego $N(\mu, 1)$ z nieznaną wartością oczekiwaną μ i wariancją 1. Na podstawie tej próbki zbudowano w standardowy sposób przedział ufności na poziomie 0.95 dla μ:

$$\left[\bar{X} - \frac{1.96}{\sqrt{n}}, \bar{X} + \frac{1.96}{\sqrt{n}}\right] = [\hat{\mu}_-, \hat{\mu}_+]$$

Pokazać, że jednostajnie najmocniejszy test hipotezy $H_0 : \mu = 3$ przeciw alternatywie $H_1 : \mu > 3$ na poziomie istotności 0.025 odrzuca H_0 wtedy i tylko wtedy, gdy $\hat{\mu}_- > 3$.
Odp. —

Zadanie 38. Niech $X_1, \dots, X_5$ będzie próbką z rozkładu prawdopodobieństwa o gęstości

$$f_\theta(x) = \begin{cases} \theta x^{-(\theta+1)}, & \text{dla } x > 1, \\ 0, & \text{poza tym.} \end{cases}$$

Rozważmy jednostajnie najmocniejszy test hipotezy $H_0 : \theta = 0.5$ przeciwko alternatywie $H_1 : \theta > 0.5$ na poziomie istotności 0.01. Wyznaczyć obszar krytyczny tego testu.
Odp. $\sum_{i=1}^{5} \ln X_i < 2.5582$

Zadanie 39. $X_1, \dots, X_{10}$ jest próbką z rozkładu normalnego o znanej wartości oczekiwanej μ i nieznanej wariancji σ^2. Rozważmy test hipotezy $H_0 : \sigma^2 \leq 4$ przeciwko alternatywie $H_1 : \sigma^2 > 4$, który jest najmocniejszy na poziomie istotności 0.05. Dla jakich wartości wariancji moc tego testu jest nie mniejsza niż 0.95? Podać zbiór $M = \{\sigma^2 : \text{ moc testu } \geq 0.95\}$.
Odp. $[18.58, +\infty)$

Zadanie 40. Każda ze zmiennych losowych $X_1, \dots, X_{100}$ ma rozkład normalny $N(\mu, \sigma^2)$ z nieznaną wartością oczekiwaną μ i znaną wariancją σ^2. Założono, że zmienne są niezależne i wyznaczono (przy tych założeniach) test jednostajnie najmocniejszy dla testowania hipotezy $H_0 : \mu = \mu_0$ przy alternatywie $H_1 : \mu > \mu_0$ na poziomie istotności 0.05. W rzeczywistości zmienne losowe $X_1, \dots, X_{100}$ mają łączny rozkład normalny, ale są skorelowane i współczynnik korelacji $Corr(X_i, X_j) = 0.1$ dla wszystkich $i \neq j$. Obliczyć faktyczny błąd pierwszego rodzaju testu z dokładnością do 0.01.
Odp. 0.31

Zadanie 41. Niech X_1, X_2, X_3, X_4 będzie próbą z rozkładu jednostajnego na przedziale $(0, \theta)$. Zakładamy, że nieznany parametr θ jest zmienną losową o rozkładzie z funkcją gęstości daną wzorem

$$\pi(\theta) = \begin{cases} \frac{4}{3}\theta^4 e^{-2\theta}, & \text{dla } \theta > 0, \\ 0, & \text{poza tym.} \end{cases}$$

Hipotezę $H_0 : \theta \leq 3$ przy alternatywie $H_1 : \theta > 3$ odrzucamy dla tych wartości $\{x_1, x_2, x_3, x_4\}$, dla których prawdopodobieństwo *a posteriori* zbioru $\{\theta : \theta > 3\}$ jest większe niż 0.5. Niech $x_{4:4} = \max\{x_1, x_2, x_3, x_4\}$. Wyznaczyć obszar krytyczny testu.
Odp. $\{(x_1, x_2, x_3, x_4) : x_{4:4} > 3 - \frac{\ln 2}{2}\}$

Zadanie 42. Niech $X_1, \dots, X_m$ będą zmiennymi losowymi o rozkładzie normalnym $N(\mu,\sigma^2)$ każda oraz niech $Y_1, \dots, Y_n$ będą zmiennymi losowymi o rozkładzie normalnym $N(\mu_2, \sigma^2)$ każda. Wszystkie zmienne są niezależne. Hipotezę $H_0 : \mu_1 = \mu_2$ przy alternatywie $H_1 : \mu_1 > \mu_2$ weryfikujemy w następujący sposób. Zliczamy liczbę S elementów w próbce $X_1, \dots, X_m$ większych od wszystkich elementów próbki $Y_1, \dots, Y_n$. Hipotezę H_0 odrzucamy, gdy $S \geq s$, gdzie s jest wartością krytyczną. Przypuśćmy, że $m = 7$ i $n = 8$. Wyznaczyć rozmiar testu, gdy $s = 2$.
Odp. 0.20

Zadanie 43. Obserwujemy n niezależnych zmiennych losowych $X_1, \ldots, X_n$ o tym samym rozkładzie o gęstości

$$f_\theta(x) = \begin{cases} \frac{2x}{\theta}, & \text{dla } x \in (0, \theta), \\ 0, & \text{poza tym,} \end{cases}$$

gdzie $\theta > 0$ jest nieznanym parametrem. Rozważmy test jednostajnie najmocniejszy dla weryfikacji hipotezy $H_0 : \theta = 1$ przy alternatywie $H_1 : \theta > 1$ na poziomie istotności 0.1. Jak najmniej liczną próbą należy dysponować, aby moc otrzymanego testu przy alternatywie $\theta = 2/3$ była nie mniejsza niż 0.9?
Odp. $n = 3$

Zadanie 44. Niech $X_1, \ldots, X_n$ będą niezależnymi zmiennymi losowymi z rozkładu normalnego $N(\mu, \sigma^2)$ z nieznanymi parametrami μ i σ^2. Rozważamy problem testowania hipotezy $H_0 : \mu = 0$ przy alternatywie $H_1 : \mu \neq 0$ za pomocą testu, który odrzuca H_0, jeśli $\frac{|\bar{X}|}{Z} > t$, gdzie $Z = \sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2}$. Dobrać stałą t tak, aby prawdopodobieństwo błędu pierwszego rodzaju testu było równe 0.05, jeśli wiadomo, że $n = 9$.
Odp. 0.632

Zadanie 45. Każda ze zmiennych losowych ma $X_1, \ldots, X_9$ rozkład normalny z nieznaną wartością oczekiwaną μ_1 i wariancją 1, a każda ze zmiennych losowych $Y_1, \ldots, Y_9$ rozkład normalny z nieznaną wartością oczekiwaną μ_2 i wariancją 4. Założono, że wszystkie zmienne losowe są niezależne i wyznaczono, przy tych założeniach, test oparty na ilorazie wiarogodności dla testowania hipotezy $H_0 : \mu_1 = \mu_2$ przy alternatywie $H_1 : \mu_1 \neq \mu_2$ na poziomie istotności 0.05. W rzeczywistości założenie to nie jest spełnione: co prawda pary zmiennych $(X_1, Y_1), \ldots, (X_9, Y_9)$ są niezależne, ale X_i, Y_i są zależne i współczynnik korelacji $Corr(X_i, Y_i) = 0.5$ dla $i = 1, \ldots, 9$. Obliczyć faktyczne prawdopodobieństwo błędu pierwszego rodzaju α i moc testu β przy alternatywie $\mu_1 = \mu_2 + 2$.
Odp. $\alpha = 0.01$, $\beta = 0.82$

Zadanie 46. Niech $X_1, \ldots, X_{15}$ będą niezależnymi zmiennymi losowymi z rozkładu normalnego $N(\mu_1, \sigma^2)$, a $Y_1, \ldots, Y_{15}$ niezależnymi zmiennymi losowymi z rozkładu normalnego $N(\mu_2, \sigma^2)$. Wszystkie zmienne są niezależne, a parametry μ_1, μ_2, σ^2 są nieznane. Testujemy hipotezę $H_0 : \mu_1 = \mu_2$ przy alternatywie $H_1 : \mu_1 \neq \mu_2$. Hipotezę H_0 odrzucamy, gdy spełniona jest nierówność

$$\frac{|\bar{X} - \bar{Y}|}{S} > c,$$

gdzie $\bar{X} = \sum_{i=1}^{15} X_i$, $\bar{Y} = \sum_{i=1}^{15} Y_i$ i $S^2 = \sum_{i=1}^{15}(X_i - Y_i)^2$. Wyznaczyć c tak, aby rozmiar testu był równy 0.05.
Odp. 0.4973

Zadanie 47. Obserwujemy niezależne zmienne losowe $X_1, X_2, X_3, Y_1, Y_2, Y_3, Y_4$. Zmienne losowe X_1, X_2, X_3 mają ten sam rozkład o dystrybuancie F_{μ_1}, a zmienne losowe Y_1, Y_2, Y_3, Y_4 mają ten sam rozkład o dystrybuancie F_{μ_2}. Dystrybuanta F_μ spełnia warunek $F_\mu = F(x - \mu)$ dla pewnej ustalonej, nieznanej, ciągłej, ściśle rosnącej dystrybuanty F. Weryfikujemy hipotezę $H_0 : \mu_1 = \mu_2$ przy alternatywie $H_1 : \mu_1 > \mu_2$ stosując test o obszarze krytycznym $K = \{S : S > 13\}$, gdzie S jest sumą rang zmiennych losowych X_1, X_2, X_3 w próbce złożonej ze wszystkich obserwacji ustawionych w ciąg rosnący. Wyznaczyć rozmiar testu.
Odp. 12/35

Zadanie 48. Na podstawie prostej próby losowej $X_1, \ldots, X_{20}$ testowano hipotezę $H_0 : \sigma^2 = 1$ przy alternatywie $H_1 : \sigma^2 > 1$, gdzie σ^2 jest parametrem odpowiadającym za wariancję zmiennej losowej X_i, za pomocą testu o obszarze krytycznym $K = \left\{\sum_{i=1}^{20} X_i^2 > t\right\}$. Wiadomo, że zmienne losowe mają rozkład zadany gęstością

$$f_\theta(x) = \theta|x|e^{-\theta x^2}, \quad \text{dla} \quad x \in R,$$

gdzie $\theta > 0$ jest nieznanym parametrem. Wyznaczyć wartość krytyczną t tak, by test miał poziom istotności 0.05.
Odp. 27.8793

Zadanie 49. Niech $X_1, \ldots, X_6$ będą niezależnymi zmiennymi losowymi o tym samym rozkładzie o gęstości

$$f_\theta(x) = \begin{cases} \theta x^{\theta-1}, & \text{dla } 0 < x < 1, \\ 0, & \text{poza tym,} \end{cases}$$

gdzie $\theta > 0$ jest nieznanym parametrem. Weryfikujemy hipotezę $H_0 : \theta = 1$ przy alternatywie $H_1 : \theta > 1$ testem jednostajnie najmocniejszym na poziomie istotności 0.05. Wyznaczyć moc tego testu przy $\theta = 3$.
Odp. 0.79

Zadanie 50. Obserwujemy X_1, X_2, X_3, X_4 niezależnych zmiennych losowych o tym samym rozkładzie Pareto o gęstości

$$f_{\theta_1}(x) = \begin{cases} \frac{2^{\theta_1}\theta_1}{(2+x)^{\theta_1+1}}, & \text{dla } x > 0, \\ 0, & \text{poza tym} \end{cases}$$

i Y_1, Y_2, Y_3, Y_4, Y_5 niezależnych zmiennych losowych o tym samym rozkładzie Pareto o gęstości

$$f_{\theta_2}(x) = \begin{cases} \frac{2^{\theta_2}\theta_2}{(2+x)^{\theta_2+1}}, & \text{dla } x > 0, \\ 0, & \text{poza tym,} \end{cases}$$

gdzie θ_1 i θ_2 są nieznanymi parametrami dodatnimi. Wszystkie zmienne losowe są niezależne. Testujemy hipotezę $H_0 : \frac{\theta_1}{\theta_2} = 1$ przy alternatywie $H_1 : \frac{\theta_1}{\theta_2} > 1$ za pomocą testu o obszarze krytycznym

$$K = \left\{ \frac{\hat{\theta}_1}{\hat{\theta}_2} > t \right\},$$

gdzie $\hat{\theta}_1$ i $\hat{\theta}_2$ są estymatorami największej wiarogodności odpowiednio parametrów θ_1 i θ_2 wyznaczonymi na podstawie prób losowych X_1, X_2, X_3, X_4 i Y_1, Y_2, Y_3, Y_4, Y_5. Dobrać stałą t tak, aby otrzymać test o rozmiarze 0.05.
Odp. 3.347

Zadanie 51. Niech $X_1, \ldots, X_{10}$ będą niezależnymi zmiennymi losowymi z rozkładu normalnego, przy czym $EX_i = 0$ i $VarX_i = \frac{\sigma^2}{i}$, gdzie σ^2 jest nieznanym parametrem. Rozważamy jednostajnie najmocniejszy test hipotezy $H_0 : \sigma^2 \leq 4$ przy alternatywie $H_1 : \sigma^2 > 4$ na poziomie istotności 0.05. Wyznaczyć zbiór S tych wartości wariancji σ^2, dla których moc tego testu jest nie mniejsza niż 0.95.
Odp. $(18.584, +\infty)$

Zadanie 52. Zakładając, że zmienne losowe $X_1, \ldots, X_5$ są niezależne i mają rozkłady normalne $X_i \sim N(\mu\sqrt{i}, 1)$ zbudowano test jednostajnie najmocniejszy dla weryfikacji hipotezy $H_0 : \mu = 0$ przy alternatywie $H_1 : \mu > 0$ na poziomie istotności 0.05. W rzeczywistości okazało się, że wektor $(X_1, \ldots, X_5)$ ma rozkład normalny taki, że $EX_i = \mu\sqrt{i}$ oraz

$$Cov(X_i, X_j) = \begin{cases} 0.5, & \text{gdy } |i-1| = 1, \\ 1, & \text{gdy } i = j, \\ 0, & \text{poza tym.} \end{cases}$$

Wyznaczyć rzeczywisty rozmiar testu.
Odp. 0.11

Zadanie 53. Obserwujemy X_1, X_2, X_3, X_4 niezależnych zmiennych losowych o tym samym rozkładzie Pareto o gęstości

$$f_{\theta_1}(x) = \begin{cases} \frac{\theta_1}{(1+x)^{\theta_1+1}}, & \text{dla } x > 0, \\ 0, & \text{poza tym} \end{cases}$$

i Y_1, Y_2, Y_3, Y_4, Y_5 niezależnych zmiennych losowych o tym samym rozkładzie Pareto o gęstości

$$f_{\theta_2}(x) = \begin{cases} \frac{\theta_2}{(1+x)^{\theta_2+1}}, & \text{dla } x > 0, \\ 0, & \text{poza tym,} \end{cases}$$

gdzie θ_1 i θ_2 są nieznanymi parametrami dodatnimi. Wszystkie zmienne losowe są niezależne. Testujemy hipotezę $H_0 : \frac{\theta_1}{\theta_2} = 2$ przy alternatywie $H_1 : \frac{\theta_1}{\theta_2} < 2$ za pomocą testu o obszarze krytycznym

$$K = \left\{ \frac{\hat{\theta}_1}{\hat{\theta}_2} < t \right\},$$

gdzie $\hat{\theta}_1$ i $\hat{\theta}_2$ są estymatorami największej wiarogodności odpowiednio parametrów θ_1 i θ_2 wyznaczonymi na podstawie prób losowych X_1, X_2, X_3, X_4 i Y_1, Y_2, Y_3, Y_4, Y_5. Dobrać stałą t tak, aby otrzymać test o rozmiarze 0.05.
Odp. 0.6511

Zadanie 54. Niech $X_1, \ldots, X_6$ będą niezależnymi zmiennymi losowymi o tym samym rozkładzie jednostajnym na przedziale $(0, \theta)$, gdzie $\theta > 0$ jest nieznanym parametrem. Zbudowano test jednostajnie najmocniejszy dla weryfikacji hipotezy $H_0 : \theta = 1$ przy alternatywie $H_1 : \theta \neq 1$ na poziomie istotności 0.125. Wyznaczyć obszar krytyczny tego testu.
Odp. $\left\{ \max\{X_1, \ldots, X_6\} \in \left(0, \frac{\sqrt{2}}{2}\right) \cup (1, +\infty) \right\}$

Zadanie 55. Obserwujemy niezależne zmienne losowe $X_1, X_2, X_3, X_4, Y_1, Y_2, Y_3, Y_4, Y_5$. Zmienne losowe X_1, X_2, X_3, X_4 mają ten sam rozkład o dystrybuancie F_{μ_1}, a zmienne losowe Y_1, Y_2, Y_3, Y_4, Y_5 mają ten sam rozkład o dystrybuancie F_{μ_2}. Dystrybuanta F_μ spełnia warunek

$$F_\mu(x) = F(x - \mu)$$

dla pewnej ustalonej, nieznanej, ciągłej, ściśle rosnącej dystrybuanty F. Weryfikujemy hipotezę $H_0 : \mu_1 = \mu_2$ przy alternatywie $H_1 : \mu_1 \neq \mu_2$ stosując test o obszarze krytycznym

$$K = \{S : S \leq 13 \vee S \geq 27\},$$

gdzie S jest sumą rang zmiennych losowych X_1, X_2, X_3, X_4 w próbce złożonej ze wszystkich obserwacji ustawionych w ciąg rosnący. Wyznaczyć rozmiar testu.
Odp. 7/63

Test chi-kwadrat

Zadanie 56. Testujemy niezależność dwóch cech z tablicy kontygencyjnej. Jedna z cech przyjmuje n, a druga m możliwych wartości. Ilość obserwacji w każdej z $(n \cdot m)$ komórek jest wystarczająca, aby zastosować test niezależności chi-kwadrat. Odpowiednia statystyka będzie miała asymptotycznie rozkład chi-kwadrat. Wyznaczyć ilość stopni swobody tego rozkładu.
Odp. $(n-1)(m-1)$

Zadanie 57. Wykonano 120 rzutów dwiema kośćiami do gry: czarną i białą. 45 razy na białej kości wypadło więcej oczek niż na czarnej, 50 razy na białej kości wypadło mniej oczek niż na czarnej oraz 25 razy na białej kości wypadła ta sama liczba oczek co na czarnej. Rozważmy hipotezę H_0 : *obie kości są rzetelne i wynik rzutu kością białą jest niezależny od wyniku rzutu kością czarną*. W oparciu o podane wyniki przeprowadzono test chi-kwadrat. Pokazać, że na poziomie istotności 0.1 test nie prowadzi do odrzucenia H_0.
Odp. —

Zadanie 58. Za pomocą testu chi-kwadrat zgodności testowano hipotezę, iż n elementowa próbka pochodzi z rozkładu Poissona o wartości oczekiwanej równej jeden. Mamy niepełną informację o próbce, na podstawie której przeprowadzono test:

k	0	1	2	3 lub więcej
Ilość obserwacji	$n-70-4-25$	70	40	25

Podać najmniejszą możliwą liczebność próbki n, jeśli wiadomo, iż na poziomie istotności 0.05 nie znaleziono podstaw do odrzucenia hipotezy o zgodności.
Odp. 196

Zadanie 59. Wykonujemy n rzutów kością do gry i weryfikujemy hipotezę H_0 mówiącą, że kość jest rzetelna. Standardowy test chi-kwadrat na poziomie istotności 0.001 odrzuca hipotezę zerową, jeśli obliczona wartość statystyki testowej przekracza 20.515. Przypuśćmy, że wykonaliśmy tylko $n = 6$ rzutów. Jest to zbyt mało, żeby asymptotyczne przybliżenie rozkładu chi-kwadrat było zadowalające. Wyznaczyć faktyczny rozmiar testu.
Odp. $1/6^5$

Zadanie 60. Zmienna losowa X przyjmuje wartości 1 lub 2 z jednakowym prawdopodobieństwem 0.5. Zmienna losowa Y przyjmuje wartości $1, 2, \ldots, k$. Dysponujemy próbką z łącznego rozkładu prawdopodobieństwa zmiennych losowych X i Y, złożoną z n par obserwacji. Niech n_{ij} oznacza liczbę takich par, dla których zmienna X przyjęła wartość i, zaś Y wartość j $(i = 1, 2; j = 1, 2, \ldots, k)$. W celu weryfikacji hipotezy o niezależności zmiennych X i Y, czyli hipotezy

$$H_0 : P(X = i, Y = j) = \frac{1}{2} P(Y = j) \text{ dla } i = 1, 2;\ j = 1, 2, \ldots, k,$$

używamy statystyki

$$\chi^2 = \sum_{i=1}^{2} \sum_{j=1}^{k} \frac{(n_{ij} - n\hat{p}_{ij})^2}{n\hat{p}_{ij}/2}, \quad \text{gdzie} \quad \hat{p}_{ij} = \frac{n_{1j} + n_{2j}}{2}.$$

Przy $n \to \infty$, rozkład tej statystyki zmierza do rozkładu chi-kwadrat. Wyznaczyć liczbę stopni swobody rozkładu granicznego.
Odp. k

Zadanie 61. Dysponujemy danymi o liczbie szkód zgłoszonych przez klientów $1, 2, \ldots, k$ w ciągu n lat. Niech $S_i(n)$ oznacza sumaryczną liczbę szkód dla klienta numer i w ciągu n lat. Wiemy, że $S_1(n), \ldots, S_k(n)$ są niezależnymi zmiennymi o rozkładzie Poissona. Mamy też pewne przypuszczenia dotyczące intensywności pojawiania się szkód, czyli wartości oczekiwanych tych zmiennych, ale nie jesteśmy ich pewni. Weryfikujemy hipotezę statystyczną

$$H_0 : \text{dla każdego } i = 1, \ldots, k \text{ zmienna losowa } S_i(n) \text{ ma rozkład Poissona z parametrem } n\lambda_i.$$

Hipotetyczne intensywności $\lambda_1, \ldots, \lambda_k$ są danymi, ustalonymi liczbami dodatnimi. Używamy pewnej odmiany testu chi-kwadrat: obliczamy statystykę

$$\chi^2 = \sum_{i=1}^{k} \frac{(S_i(n) - n\lambda_i)^2}{n\lambda_i}.$$

Jaki jest rozkład graniczny tej statystyki, jeśli H_0 jest prawdziwa i $n \to \infty$?
Odp. rozkład chi-kwadrat z k stopniami swobody

Zadanie 62. Wykonujemy n rzutów kością do gry i weryfikujemy hipotezę H_0 mówiącą, że kość jest rzetelna, tzn. że każda liczba oczek pojawia się z jednakowym prawdopodobieństwem $\frac{1}{6}$. Standardowy test chi-kwadrat na poziomie istotności 0.01 odrzuca hipotezę zerową, jeżeli wartość statystyki chi-kwadrat przekracza 15.0863 (kwantyl rzędu 0.99 rozkładu chi-kwadrat z pięcioma stopniami swobody). Przypuśćmy, że wykonaliśmy tylko $n = 6$ rzutów. Jest to zbyt mało, żeby asymptotyczne przybliżenie rozkładu chi-kwadrat było zadowalające. Wyznaczyć faktyczny rozmiar testu odrzucamy H_0, jeśli wartość statystyki chi-kwadrat przekroczy 15.0863.
Odp. $\frac{31}{6^5}$

Zadanie 63. W pewnym portfelu ubezpieczeń samochodowych liczącym 450 ubezpieczonych odnotowano, że w ciągu roku ubezpieczenia 40 ubezpieczonych kobiet uczestniczyło w wypadku, 30 ubezpieczonych mężczyzn uczestniczyło w wypadku, 160 ubezpieczonych kobiet nie uczestniczyło w wypadku oraz 220 ubezpieczonych mężczyzn nie uczestniczyło w wypadku. Ile wynosi wartość statystyki testu chi-kwadrat niezależności i czy (na poziomie istotności 0.05) istnieją podstawy do odrzucenia hipotezy o niezależności wystąpienia wypadku od płci ubezpieczonego?
Odp. 5.41; należy odrzucić hipotezę o niezależności

Inne zagadnienia testowania

Zadanie 64. Mamy trzy niezależne, dziesięcioelementowe próbki proste pobrane z trzech populacji normalnych: $X_{i,1}, \ldots, X_{i,10} \sim N(\mu_i, \sigma^2)$, $i = 1, 2, 3$ o tej samej (nieznanej) wariancji σ^2. W każdym z trzech przypadków policzono średnią $\bar{X}_i = \frac{1}{10}\sum_{i=1}^{10} X_{i,j}$ oraz wariancję z próbki $S_i^2 = \frac{1}{9}\sum_{i=1}^{10}(X_{i,j} - \bar{X}_i)^2$. Uzyskano następujące wyniki:

i	1	2	3
$\bar{X}_i$	30	31	32
S_i^2	15/9	25/9	20/9

Przeprowadzono testy F analizy wariancji na poziomie istotności 0.05 dla weryfikacji każdej z następujących hipotez: $H_{12} : \mu_1 = \mu_2$, $H_{23} : \mu_2 = \mu_3$, $H_{13} : \mu_1 = \mu_3$, $H_{123} : \mu_1 = \mu_2 = \mu_3$. Które z hipotez zostaną odrzucone, a które nie?
Odp. odrzucone H_{123} i H_{13}

Zadanie 65. Rozważamy model K obiektów obserwowanych przez T okresów czasu, gdzie zarówno K jak i T są dużymi liczbami. Przyjmujemy następujące założenia:
• dla każdego $k = 1, \ldots, K$ oraz $t = 1, \ldots, T$ warunkowy rozkład zmiennej losowej $X_{t,k}$ przy danej wartości zmiennej μ_k, jest rozkładem normalnym o wartości oczekiwanej i wariancji (μ_k, σ^2);
• dla każdego $k = 1, \ldots, K$ rozkład zmiennej losowej μ_k jest rozkładem normalnym o wartości oczekiwanej i wariancji (μ, a^2).
Przyjmijmy typowe oznaczenia dla średnich obiektowych i średniej ogólnej:

$$\bar{X}_k = \frac{1}{T}\sum_{t=1}^{T} X_{t,k}, \ k = 1, \ldots, K, \quad \text{oraz} \quad \bar{X} = \frac{1}{K}\sum_{k=1}^{K} \bar{X}_k.$$

Międzyobiektową i wewnątrzobiektową sumę kwadratów odchyleń oznaczmy:

$$SSB = \sum_{k=1}^{K}(\bar{X}_k - \bar{X})^2, \quad SSW = \sum_{t=1}^{T}\sum_{k=1}^{K}(X_{t,k} - \bar{X}_k)^2.$$

Wiadomo, że zmienne losowe SSB i SSW są niezależne,

$$E(SSB) = (K-1)\left(a^2 + \frac{\sigma^2}{T}\right), \quad E(SSW) = K(T-1)\sigma^2.$$

Dobrać stałą $const$ tak, aby wartość oczekiwana wyrażenia $const \cdot \frac{SSW}{SSB}$ wyniosła $\frac{\sigma^2}{a^2T+\sigma^2}$.
Odp. $\frac{K-3}{TK(T-1)}$

Zadanie 66. Rozpatrzmy standardowy model jednokierunkowej analizy wariancji. Niech X_{ij} będą niezależnymi zmiennymi losowymi o rozkładach normalnych ($i = 1, \ldots, k$, $j = 1, \ldots, n_i$), przy czym $EX_{ij} = \mu_i$ i $VarX_{ij} = \sigma^2$. Przyjmijmy typowe oznaczenia:

$$SSW = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2, \quad SST = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2,$$

gdzie

$$\bar{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij} \quad \bar{X} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}, \quad n = \sum_{i=1}^{k} n_i.$$

Przy założeniu, że hipoteza o jednorodności jest prawdziwa, czyli że $\mu_1 = \cdots = \mu_k$, obliczyć $E\left(\frac{SSW}{SST}\right)$.
Odp. $\frac{n-k}{n-1}$