Wstęp do statystyki matematycznej
Lista 11

1. Niech $X = (X_1, ..., X_n)$ będzie próbą z rozkładu wykładniczego $\mathcal{E}(\lambda)$. Uzasadnić, że $Q(X,\lambda) = \frac{2n\overline{X}}{\lambda}$ jest funkcją centralną. Skonstruować przedział ufności dla parametru λ na poziomie ufności $1-\alpha$. Dla $n = 10, \alpha = 0.04$ podać postać końców tego przedziału.

2. Niech $X_1, ..., X_n$ będzie próbą z rozkładu $\mathcal{N}(\theta, \theta)$. Znaleźć funkcję centralną i skonstruować przedział ufności dla θ na poziomie ufności $1-\alpha$.

3. Niech $X_1, ..., X_n$ będzie próbą z rozkładu jednostajnego na odcinku $[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$, $\theta \in R$. Skonstruować przedział ufności dla parametru θ na poziomie ufności $1-\alpha$.

4. Niech $X = (X_1, ..., X_m)$ i $Y = (Y_1, ..., Y_n)$ będą niezależnymi próbami z rozkładów $\mathcal{N}(\mu_1, \sigma_1^2)$, $\mathcal{N}(\mu_2, \sigma_2^2)$. Załóżmy, że $\sigma_1 = \sigma_2 = \sigma$ i oznaczmy $\theta = (\mu_1, \mu_2, \sigma)$ i niech $g(\theta) = \mu_1 - \mu_2$ będzie funkcją parametryczną. Uzasadnić, że
$$Q(X, Y, g(\theta)) = \frac{\overline{X} - \overline{Y} - g(\theta)}{\sqrt{\sum_i (X_i - \overline{X})^2 + \sum_j (Y_j - \overline{Y})^2}} \sqrt{\frac{m+n}{mn(m+n-2)}}$$
jest funkcją centralną dla $g(\theta)$. Skonstruować przedział ufności dla $g(\theta)$ na poziomie ufności $1-\alpha$.

5. Dla modelu z zad. 4, ale przy dowolnych σ_1, σ_2 i $\theta = (\mu_1, \mu_2, \sigma_1, \sigma_2)$ oraz $g(\theta) = \sigma_1^2/\sigma_2^2$ skonstruować przedział ufności dla $g(\theta)$ na poziomie ufności $1-\alpha$.

6. Niech $X_1, ..., X_n$ będzie próbą z rozkładu $\mathcal{N}(\mu, \sigma^2)$, gdzie $\sigma > 0$ jest znane. Podać postać przedziału ufności dla parametru μ na poziomie ufności $1-\alpha$ i wykazać, że najkrótszy przedział otrzymuje się, przyjmując $c_1 = \Phi^{-1}(\alpha/2)$, $c_2 = \Phi^{-1}(1-\alpha/2)$.

7. Niech Y_n bedzie zmienną losową o rozkładzie chi-kwadrat z n stopniami swobody. Wykazać, że a) $\frac{Y_n - n}{\sqrt{2n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$; b) $\sqrt{2Y_n} - \sqrt{2n-1} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$; c) $\sqrt{\frac{9n}{2}}\left(\sqrt[3]{\frac{Y_n}{n}} - 1\right) - \frac{\sqrt{2n}}{3(n-1)} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$ (Wilson-Hilferty, 1931). Ciąg zmiennych losowych w c) daje najlepsze przybliżenie rozkładu chi-kwadrat. Wykorzystać powyższe fakty do określenia asymptotycznych funkcji centralnych i konstrukcji asymptotycznych przedziałów ufności dla $g(\theta) = \sigma^2$ na poziomie ufności $1-\alpha$ w modelu określonym w zad. 6, ale z nieznanymi μ i σ oraz $\theta = (\mu, \sigma)$.

8. Niech $X_1, ..., X_n$ będzie próbą z rozkładu Poissona z parametrem λ. Udowodnić, że a) $\sqrt{n}(\overline{X} - \lambda)/\sqrt{\overline{X}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$; b) $2\sqrt{n}(\sqrt{\overline{X}} - \sqrt{\lambda}) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$. Wykorzystać powyższe fakty do określenia asymptotycznych funkcji centralnych i konstrukcji asymptotycznych przedziałów ufności dla parametru λ na poziomie ufności $1-\alpha$. Zauważyć, że oba przedziały mają tę samą długość i są tylko przesunięte względem siebie.

9. Rozważmy model dwumianowy o gęstości $p_\theta(x) = \begin{pmatrix} n \\ x \end{pmatrix} \theta^x (1-\theta)^{n-x} \mathbf{1}_{\{0,1,\ldots,n\}}(x)$, $\theta \in (0,1)$, względem miary liczącej. Skonstruować przedział ufności dla parametru θ na poziomie ufności $1-\alpha$ w oparciu o nierówność: a) Hoeffdinga $P(|X/n - \theta| \geqslant \epsilon) \leqslant 2\exp\{-2n\epsilon^2\}$; b) Bernsteina $P(|X/n - \theta| \geqslant \epsilon) \leqslant 2\exp\{-n\epsilon^2/2(\theta + \epsilon)\}$.

Wstęp do statystyki matematycznej
Lista 12

1. Rozważmy model dwumianowy jak w zad. 9 poprzedniej listy, gdzie $n = 15$. Skonstruować test UMP dla testowania hipotezy $H_0 : \theta \leqslant 0.2$ lub $\theta \geqslant 0.7$ przeciwko $H_1 : 0.2 < \theta < 0.7$ na poziomie istotności $\alpha = 0.1$. Wyznaczyć moc tego testu dla alternatywy $\theta = 0.5$.

2. Udowodnić, że w jednoparametrowej naturalnej rodzinie wykładniczej nie istnieje test UMP hipotezy $H_0 : \vartheta_1 \leqslant \vartheta \leqslant \vartheta_2$ przeciwko $H_1 : \vartheta < \vartheta_1$ lub $\vartheta > \vartheta_2$.

3. Niech $X_1, ..., X_n$ będzie próbą z rozkładu wykładniczego z parametrem λ. Rozważmy hipotezę $H_0 : \lambda = \lambda_0$ przeciwko $H_1 : \lambda \neq \lambda_0$ oraz statystykę $T = \dfrac{n\overline{X}}{\lambda_0}$ (por. zad. 1 lista 11). Pokazać, że T jest statystyką testową testu UMPU, który odrzuca H_0 na poziomie α, gdy $T < t_{1\alpha}$ lub $T > t_{2\alpha}$. Podać warunki określające $t_{1\alpha}$ i $t_{2\alpha}$, korzystając z twierdzenia o postaci testu UMPU. Dla $n = 10$ i $\alpha = 0.05$ obliczyć $t_{1\alpha}$, $t_{2\alpha}$ za pomocą tablic rozkładu chi-kwadrat.

4. Dla modelu z zad. 1 i $n = 10$ skonstruować test UMPU dla hipotezy $H_0 : \theta = 0.3$ przeciwko $H_1 : \theta \neq 0.3$ na poziomie α. Obliczyć moc tego testu dla alternatywy $\theta = 0.4$.

5. Niech $(X_1, Y_1), ..., (X_n, Y_n)$ będzie próbą z dwuwymiarowego rozkładu normalnego $\mathcal{N}(m, \Sigma)$, gdzie $m \in R^2$, a
$$\Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma\tau \\ \rho\sigma\tau & \tau^2 \end{pmatrix}, \ \sigma, \tau > 0, \ \rho \in (-1, 1).$$
Niech $T = \dfrac{R}{\sqrt{1 - R^2}}\sqrt{n-2}$, gdzie
$$R = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})/S_X S_Y$$
jest współczynnikiem korelacji z próby. Udowodnić, że test UMPU dla hipotezy $H_0 : \rho = 0$ przeciwko $H_1 : \rho \neq 0$ na poziomie α, przy parametrach zakłócających m, σ, τ, odrzuca hipotezę H_0, gdy $|T| > c$. Udowodnić, że przy H_0 statystyka T ma rozkład Studenta z $n - 2$ stopniami swobody.

6. (rozszerzenie zad. 3) Niech $X_1, ..., X_n$ będzie próbą z rozkładu wykładniczego z parametrami θ, λ (por. lista 4, zad. 10g). Niech $T = 2\sum_{i=1}^{n}(X_i - X_{(1)})$. Pokazać, że test UMPU hipotezy $H_0 : \lambda = 1$ przeciwko $H_1 : \lambda \neq 1$ odrzuca H_0 na poziomie α, przy parametrze zakłócającym θ, gdy $T < c_1$ lub $T > c_2$, gdzie c_1, c_2 sa wyznaczone przez warunek $P(\chi^2_{2n-2} \in [c_1, c_2]) = P(\chi^2_{2n} \in [c_1, c_2]) = 1 - \alpha$.