Zadanie 1. Niech $X_1, \ldots, X_8$ będzie próbą z rozkładu normalnego z wartością oczekiwaną θ i wariancją 1. Nieznany parametr θ jest z kolei zmienną losową o rozkładzie normalnym z wartością oczekiwaną 0 i wariancją 1. Wyznaczyć bayesowski przedział ufności dla parametru θ, to znaczy przedział $[a, b]$, gdzie $a = a(X_1, \ldots, X_8)$, $b = b(X_1, \ldots, X_8)$, taki, że $P(\theta < a | X_1, \ldots, X_8) = 0.05 = P(\theta > b | X_1, \ldots, X_8)$.
Odp. $[\bar{X} - 0.427, \bar{X} + 0.427]$

Zadanie 2. $X_1, \ldots, X_{10}$ jest próbą z rozkładu normalnego $N(\mu, \sigma^2)$, gdzie μ i σ^2 są nieznanymi parametrami. Niech $[L, U]$ będzie przedziałem ufności dla parametru μ taki, że dla wszystkich μ i σ^2

$$P_{\mu,\sigma^2}(L > \mu) = P_{\mu,\sigma^2}(U < \mu) = 0.025.$$

Niech $(-\infty, W]$ będzie jednostronnym przedziałem ufności dla μ takim, że $P_{\mu,\sigma^2}(W < \mu) = 0.01$ dla wszystkich μ i σ^2. Oba przedziały ufności zbudowane są w standardowy sposób w oparciu o średnią i wariancję $\bar{X}$ i S^2 z próbki. Wiadomo, że $L = -0.262$ i $U = 4.262$. Wyznaczyć W.
Odp. 4.821

Zadanie 3. Niech $X_1, \ldots, X_8$ będzie próbą z rozkładu jednostajnego na przedziale $(0, \theta)$, gdzie $\theta > 0$ jest nieznanym parametrem. Znaleźć najmniejsze c takie, żeby przedział $[\max\{X_1, \ldots, X_8\}, c \max\{X_1, \ldots, X_8\}]$ był przedziałem ufności dla θ na poziomie ufności 0.9375.
Odp. 1.4142

Zadanie 4. Zakładamy, że X_1, X_2, X_3, X_4, X_5 jest próbą losową z rozkładu 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. Skonstruować przedział ufności $[\underline{\theta}, \overline{\theta}]$ dla parametru θ (na poziomie $1 - \alpha = 0.90$) tak, żeby $P_\theta(\underline{\theta} < \theta) = 0.05 = P_\theta(\overline{\theta} > \theta)$.
Odp. $[\frac{3.4}{2S}, \frac{18.31}{2S}]$, gdzie $S = -\sum_{i=1}^{5} \ln X_i$

Zadanie 5. Niech $X_1, \ldots, X_n$ będzie próbką n niezależnych realizacji zmiennej losowej X o rozkładzie ciągłym i niech $X_1^{(n)}, \ldots, X_n^{(n)}$ oznacza uporządkowaną (od wartości najmniejszych do największych) próbę. Budujemy przedział ufności dla mediany zmiennej X o postaci $U_n = (X_2^{(n)}, X_{n-1}^{(n)})$. Oznaczmy przez n^* najmniejszą z tych wartości n, dla których prawdopodobieństwo pokrycia mediany przez przedział U_n przekracza 0.95. Wyznaczyć n^*.
Odp. 9

Zadanie 6. Dwie niezależne próbki proste $X_1, \ldots, X_n$ i $Y_1, \ldots, Y_n$ pochodzą z tego samego rozkładu normalnego o parametrach μ i σ^2. Jeden statystyk ma do dyspozycji pierwszą próbkę, drugi zaś drugą. Obaj statystycy znają wariancję σ^2, żaden nie zna wartości oczekiwanej μ. Każdy z nich buduje na podstawie swojej próbki przedział ufności dla μ na poziomie ufności 0.8. Wyznaczyć prawdopodobieństwo, że zbudowane przedziały okażą się rozłączne.
Odp. 0.07

Zadanie 7. Na podstawie próbki $X_1, \ldots, X_n$ z rozkładu normalnego $N(\mu, \sigma^2)$ z nieznanymi parametrami μ i σ^2 budujemy przedział ufności $(\underline{\sigma^2}, \overline{\sigma^2})$ dla wariancji na poziomie ufności 0.95. Metodę wybieramy możliwie prostą, korzystając na przykład z przybliżenia rozkładu chi-kwadrat z k stopniami swobody rozkładem normalnym $N(k, 2k)$. Względny błąd estymacji przedziałowej mierzymy za pomocą ilorazu $R = \frac{\overline{\sigma^2} - \underline{\sigma^2}}{2\sigma^2}$. Dla jakiego rozmiaru n próbki $E(R) \approx 0.01$.
Odp. 75000

Zadanie 8. Jacek i Placek dostaną próbkę prostą $X_1, \ldots, X_{10}$ z rozkładu normalnego $N(\mu, \sigma^2)$. Obaj nie znają wartości oczekiwanej μ, ale Jacek zna wariancję σ^2, a Placek jej nie zna. Obaj budują w standardowy sposób przedziały ufności dla μ na poziomie ufności 0.95. Placek się chwali: „mam szansę 10%, że mój przedział ufności będzie przynajmniej x razy krótszy, niż Twój". Znaleźć x.
Odp. $x \approx 1.25$

Zadanie 9. Niech X_1, X_2, X_3, X_4, X_5 będzie próbką z rozkładu wykładniczego o gęstości prawdopodobieństwa:

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

Parametr θ jest nieznany. Wiadomo, że estymatorem największej wiarogodności tego parametru jest $\hat{\theta} = 5/S_5$, gdzie $S_5 = X_1 + X_2 + X_3 + X_4 + X_5$. Należy zbudować przedział ufności dla parametru θ postaci

$$[\underline{\theta}, \overline{\theta}] = \left[\frac{\underline{a}}{S_5}, \frac{\overline{a}}{S_5}\right].$$

Żądamy, żeby ten przedział był symetryczny w tym sensie, że $P_\theta(\theta < \underline{\theta}) = P_\theta(\theta > \overline{\theta})$. Wyznaczyć stałe $\underline{a}$ i $\overline{a}$ tak, żeby otrzymać przedział na poziomie ufności 0.95.
Odp. $\underline{a} = 1.62$. $\overline{a} = 10.24$

Zadanie 10. Niech $X_1, \ldots, X_{20}$ będzie próbką losową z rozkładu normalnego $N(\mu, \sigma^2)$, z nieznanymi parametrami μ i σ^2. Niech

$$\bar{X}_{10} = \frac{1}{10}\sum_{i=1}^{10} X_i, \ \bar{X}_{20} = \frac{1}{20}\sum_{i=1}^{20} X_i, \ S^2 = S_{10}^2 = \frac{1}{9}\sum_{i=1}^{10}(X_i - \bar{X}_{10})^2.$$

Należy skonstruować przedział $[\bar{X}_{10} - aS, \bar{X}_{10} + aS]$ taki, że $P_{\mu,\sigma^2}(\bar{X}_{20} \in [\bar{X}_{10} - aS, \bar{X}_{10} + aS]) = 0.95$. Wyznaczyć odpowiednią liczbę a.
Odp. $2.2622/\sqrt{20}$

Zadanie 11. Załóżmy, że $X_1, \ldots, X_9$ jest próbką z rozkładu normalnego $N(\mu, \sigma^2)$ z nieznanymi parametrami μ i σ^2. Pan Ixiński miał podać przedział ufności dla μ na poziomie 0.95, ale nie znalazł tablic rozkładu t. Ponieważ miał tablice rozkładu normalnego i chi-kwadrat, więc poradził sobie tak:
1. obliczył w standardowy sposób jednostronny przedział ufności $[0, \overline{\sigma}^2]$ dla wariancji, na poziomie 0.95;
2. przyjął, że $\left[\bar{X} - \frac{1.96\overline{\sigma}}{\sqrt{9}}, \bar{X} + \frac{1.96\overline{\sigma}}{\sqrt{9}}\right]$ jest potrzebnym przedziałem dla wartości oczekiwanej, gdzie $\overline{\sigma}$ zostało wyznaczone w punkcie 1.
Obliczyć faktyczny poziom ufności takiego przedziału ufności dla μ.
Odp. ≈ 0.99

Zadanie 12. Mamy pięć niezależnych próbek z tego samego rozkładu normalnego $N(\mu, \sigma^2)$ z nieznaną wartością oczekiwaną μ i znaną wariancją σ^2, przy tym każda z tych próbek ma tę samą liczebność n. Dla każdej z pięciu próbek oddzielnie wyznaczamy w standardowy sposób przedział ufności dla μ. Niech $\left[\bar{X}_i - \frac{1.15035\sigma}{\sqrt{n}}, \bar{X}_i + \frac{1.15035\sigma}{\sqrt{n}}\right]$ będzie przedziałem obliczonym na podstawie i-tej próbki (liczba 1.15035 jest kwantylem rzędu 0.875 standardowego rozkładu normalnego). Następnie, przedział ufności oparty na wszystkich $5n$ obserwacjach wyznaczamy w sposób niestandardowy: za środek przedziału wybieramy medianę $m = med(\bar{X}_1, \bar{X}_2, \bar{X}_3, \bar{X}_4, \bar{X}_5)$. Obliczyć (z dokładnością do 0.01)

$$P_\mu\left(m - \frac{1.15035\sigma}{\sqrt{n}}, m + \frac{1.15035\sigma}{\sqrt{n}}\right).$$

Odp. 0.97

Zadanie 13. Załóżmy, że $X_1, \ldots, X_6$ jest próbą z rozkładu normalnego $N(\mu, \sigma^2)$ z nieznanymi parametrami. Zadanie polega na zbudowaniu przedziału ufności dla wariancji σ^2. Żądany poziom ufności jest równy $1-\alpha = 0.99$. Rozpatrzmy dwie metody:
Metoda S jest standardowa: budujemy przedział postaci $[0, G_S]$, gdzie $G_S = \frac{5S^2}{c}$, (S^2 oznacza nieobciążony estymator wariancji, zaś c jest odpowiednim kwantylem rozkładu chi-kwadrat);
Metoda N polega na podziale próbki na dwie części. Podpróbkę X_1, X_2, X_3 wykorzystujemy do zbudowania przedziału ufności $[0, G_{123}]$, zaś podpróbkę X_4, X_5, X_6 do zbudowania przedziału $[0, G_{456}]$. Oba te przedziały obliczamy niezależnie w standardowy sposób przyjmując poziom ufności $1 - \sqrt{\alpha} = 0.90$. Ostatecznie naszym przedziałem ufności jest $[0, G_N]$, gdzie $G_N = \max\{G_{123}, G_{456}\}$.
Porównać średnie długości przedziałów otrzymanych obiema metodami.
Odp. $EG_N = 1.58EG_S$

Zadanie 14. Każda ze zmiennych losowych $X_1, \ldots, X_{100}$ ma rozkład normalny $N(\mu, \sigma^2)$ z nieznaną wartością oczekiwaną μ i znaną wariancją σ^2. Założono, że zmienne są niezależne i w standardowy sposób zbudowano dla μ przedział ufności na poziomie ufności 0.95:

$$\left[\bar{X} - \frac{1.96\sigma}{\sqrt{10}}, \bar{X} + \frac{1.96\sigma}{\sqrt{10}}\right].$$

W rzeczywistości zmienne $X_1, \ldots, X_{100}$ mają łączny rozkład normalny, ale są skorelowane, $Corr(X_i, X_j) = 0.1$ dla wszystkich $i \neq j$. Obliczyć faktyczny poziom ufności.
Odp. 0.45

Zadanie 15. Próbka $X_1, \ldots, X_{14}$ pochodzi z rozkładu normalnego $N(\mu, \sigma^2)$ z nieznaną wartością oczekiwaną μ i nieznaną wariancją σ^2. Na podstawie tej próbki zbudowano dla μ w standardowy sposób przedział ufności na poziomie ufności $1 - \alpha = 0.995$

$$\left[\bar{X}_{14} - \frac{S \cdot t}{\sqrt{14}}, \bar{X}_{14} + \frac{S \cdot t}{\sqrt{14}}\right].$$

Niech X_{15} będzie zmienną losową pochodzącą z tego samego rozkładu, niezależną od $X_1, \ldots, X_{14}$. Obliczyć prawdopodobieństwo tego, że X_{15} należy do uprzednio wyznaczonego przedziału ufności:

$$P_{\mu,\sigma^2}\left(\bar{X}_{14} - \frac{S \cdot t}{\sqrt{14}} \leq X_{15} \leq \bar{X}_{14} + \frac{S \cdot t}{\sqrt{14}}\right).$$

Odp. 0.60

Zadanie 16. Załóżmy, że $X_1, \ldots, X_4$ jest próbką z rozkładu normalnego $N(\mu, 1)$, zaś $Y_1, \ldots, Y_9$ jest próbką z rozkładu normalnego $N(\mu, 4)$, μ jest nieznanym parametrem. Niech $\bar{X} = \frac{1}{4}\sum_{i=1}^4 X_i$, $\bar{Y} = \frac{1}{9}\sum_{i=1}^9 Y_i$. Znaleźć takie liczby r i d , żeby przedział

$$\left[r\bar{X} + (1-r)\bar{Y} - d, r\bar{X} + (1-r)\bar{Y} + d\right]$$

był przedziałem ufności dla μ na poziomie ufności $1 - \alpha = 0.95$ i przy tym długość tego przedziału ($2d$) była najmniejsza.
Odp. $r = 0.64$, $d = 0.784$

Zadanie 17. Załóżmy, że zmienne losowe $X_1, \ldots, X_{10}$ są niezależne i mają rozkłady normalne. Zmienna X_i ma rozkład $N\left(\mu, \frac{1}{i}\right)$, innymi słowy $EX_i = \mu$, $VarX_i = \frac{1}{i}$ dla $i = 1, \ldots, 10$. Wartość oczekiwana μ (jednakowa dla wszystkich zmiennych) jest nieznana. Należy zbudować przedział ufności dla μ na poziomie $1 - \alpha = 0.95$. Przedział ma być postaci $[\hat{\mu} - d, \hat{\mu} + d]$, gdzie $\hat{\mu}$ jest estymatorem największej wiarogodności parametru μ. Podać liczbę d taką, że

$$P_\mu\left(\hat{\mu} - d < \mu < \hat{\mu} + d\right) = 0.95.$$

Odp. 0.3354

Zadanie 18. Rozważmy próbkę $X_1, \ldots, X_n$ z rozkładu jednostajnego na odcinku $[0, \theta]$ (z nieznanym prawym końcem θ). Niech $M = \max\{X_1, \ldots, X_n\}$. Należy zbudować przedział ufności dla θ na poziomie 90%. Chcemy, żeby ten przedział był postaci $[aM, bM]$, gdzie liczby a i b są tak dobrane, żeby

$$P_\theta(\theta < aM) = P_\theta(\theta > bM) = 0.05.$$

Podać długość tego przedziału.
Odp. $\left(\sqrt[n]{20} - \sqrt[n]{\frac{20}{19}}\right) M$

Zadanie 19. Załóżmy, że $X_1, \ldots, X_4$ jest próbką z rozkładu normalnego $N(\mu, \sigma^2)$ o nieznanej wartości oczekiwanej i nieznanej wariancji, zaś jest X_5 zmienną losową z tego samego rozkładu, niezależną od próbki. Interpretujemy zmienną X_5 jako kolejną obserwację, która pojawi się w przyszłości, ale obecnie jest nieznana. Zbudować „przedział ufności”

$$[L, U] = [L(X_1, \ldots, X_4), U(X_1, \ldots, X_4)]$$

oparty na próbce $X_1, \ldots, X_4$ taki, że

$$P_{\mu,\sigma^2}(L(X_1, \ldots, X_4) \leq X_5 \leq U(X_1, \ldots, X_4)) = 0.95,$$

przy tym żądamy, żeby przedział był symetryczny, tzn. $\frac{1}{2}(L + U) = \bar{X}$. Używamy tutaj oznaczeń:

$$\bar{X} = \frac{1}{4}\sum_{i=1}^{4} X_i, \quad S^2 = \frac{1}{3}\sum_{i=1}^{4}(X_i - \bar{X})^2.$$

Odp. $L = \bar{X} - 3.558S$, $U = \bar{X} + 3.558S$

Zadanie 20. Zakładamy, że $X_1, \ldots, X_{10}$ są niezależnymi zmiennymi losowymi o rozkładach normalnych, przy czym $EX_i = \mu$ i $VarX_i = \frac{\sigma^2}{w_i}$. Parametry μ i σ^2 są nieznane, a wagi w_i $(i = 1, \ldots, 10)$ są znane. Zbudować przedział ufności $[\hat{\sigma}_1^2, \hat{\sigma}_2^2]$ dla σ^2 na poziomie ufności $1 - \alpha = 0.9$.
Odp. $\left[\frac{\sum_{i=1}^{10} w_i(X_i - \bar{X}_w)^2}{16.9190}, \frac{\sum_{i=1}^{10} w_i(X_i - \bar{X}_w)^2}{3.3251}\right]$, gdzie $\bar{X}_w = \frac{\sum_{i=1}^{10} w_i X_i}{\sum_{i=1}^{10} w_i}$

Zadanie 21. Losujemy n $(n \geq 3)$ niezależnych realizacji zmiennej losowej z rozkładu jednostajnego na przedziale $(0, \theta)$. Po uporządkowaniu zaobserwowanych wartości w ciąg rosnący $\{x_1, \ldots, x_n\}$ tworzymy przedział $(2x_1, 2x_{n-1})$. Dobrać najmniejsze n, przy którym prawdopodobieństwo tego, że tak utworzony przedział pokrywa wartość parametru θ jest większe niż 0.9.
Odp. 7

Zadanie 22. Niech $X_1, \ldots, X_n$ będą zmiennymi losowymi o rozkładzie Pareto$(1, a_1)$, a $Y_1, \ldots, Y_m$ będą zmiennymi losowymi o rozkładzie Pareto$(1, a_2)$, gdzie $a_1, a_2 > 0$ są nieznanymi parametrami. Wszystkie zmienne są niezależne. Na poziomie ufności $1 - \alpha$ budujemy przedział ufności $[dT, cT]$ dla ilorazu parametrów $\frac{a_1}{a_2}$ na podstawie estymatora największej wiarogodności T tego ilorazu w ten sposób, że

$$P_{a_1,a_2}\left(cT < \frac{a_1}{a_2}\right) = P_{a_1,a_2}\left(dT > \frac{a_1}{a_2}\right) = \frac{\alpha}{2}.$$

Wyznaczyć długość przedziału ufności, gdy $\alpha = 0.1$, $m = 4$, $n = 5$.
Odp. $3.02T$

Zadanie 23. Niech $(X_1, Y_1), \ldots, (X_n, Y_n)$ będą niezależnymi zmiennymi losowymi o tym samym rozkładzie normalnym z następującymi parametrami: nieznaną wartością oczekiwaną $EX_i = EY_i = \mu$, wariancją $VarX_i = \frac{1}{4}$, $VarY_i = 1$ i współczynnikiem korelacji $Corr(X_i, Y_i) = 0.5$. Osobno na podstawie prób losowych $X_1, \ldots, X_n$ i $Y_1, \ldots, Y_n$ zbudowano dwa przedziały ufności dla wartości oczekiwanej μ, każdy na poziomie ufności 0.8. Obliczyć prawdopodobieństwo, że tak zbudowane przedziały okażą się rozłączne.
Odp. 0.03

Zadanie 24. Zakładamy, że $X_1, \ldots, X_{12}$ są niezależnymi zmiennymi losowymi o rozkładach normalnych, przy czym $EX_i = \mu$ i $VarX_i = \frac{\sigma^2}{i}$, gdzie parametry $\mu \in R$ i $\sigma^2 > 0$ są nieznane. Budujemy przedział ufności $[\hat{\sigma}_1^2, \hat{\sigma}_2^2]$ dla parametru σ^2 na poziomie ufności 0.9. Niech $\bar{X} = \frac{1}{12} \sum_{i=1}^{12} X_i$ i $\bar{X}_w = \frac{1}{78} \sum_{i=1}^{12} iX_i$. Wyznaczyć $\hat{\sigma}_1^2$ i $\hat{\sigma}_2^2$ tak, by $P_{\mu,\sigma^2}(\hat{\sigma}_1^2 > \sigma^2) = P_{\mu,\sigma^2}(\hat{\sigma}_2^2 < \sigma^2) = 0.05$.
Odp. $\frac{\sum_{i=1}^{12} i(X_i - \bar{X}_w)^2}{19.6752}$ i $\frac{\sum_{i=1}^{12} i(X_i - \bar{X}_w)^2}{4.5748}$

Zadanie 25. Dysponując pięcioma niezależnymi próbkami losowymi o tej samej liczebności n, z tego samego rozkładu normalnego $N(\mu, \sigma^2)$ z nieznaną wartością oczekiwaną μ i znaną wariancją σ^2, zbudowano pięć standardowych przedziałów ufności dla parametru μ postaci $\left[\bar{X}_i - 1.2816\frac{\sigma}{\sqrt{n}}, \bar{X}_i + 1.2816\frac{\sigma}{\sqrt{n}}\right]$, gdzie $\bar{X}_i$ jest średnią z obserwacji w i-tej próbce, $i = 1, 2, 3, 4, 5$. Następnie zbudowano przedział ufności dla parametru μ postaci

$$\left[m - 1.2816\frac{\sigma}{\sqrt{n}}, m + 1.2816\frac{\sigma}{\sqrt{n}}\right],$$

gdzie $m = med\{X_1, X_2, X_3, X_4, X_5\}$. Wyznaczyć

$$P_\mu\left[m - 1.2816\frac{\sigma}{\sqrt{n}} \le \mu \le m + 1.2816\frac{\sigma}{\sqrt{n}}\right].$$

Odp. 0.98288

Zadanie 26. Niech $X_1, \ldots, X_n$ będą niezależnymi zmiennymi losowymi z rozkładu Weibulla o gęstości

$$f_\theta(x) = \begin{cases} 3\theta x^2 \exp(-\theta x^3), & \text{dla } x > 0, \\ 0, & \text{poza tym,} \end{cases}$$

gdzie $\theta > 0$ jest nieznanym parametrem. Przedział ufności dla parametru θ w oparciu o estymator największej wiarogodności $\hat{\theta}_n = \hat{\theta}_n(X_1, \ldots, X_n)$ parametru θ otrzymujemy rozwiązując nierówność

$$\left|\frac{\hat{\theta}_n - \theta}{\sigma(\theta)}\sqrt{n}\right| \le z,$$

gdzie $\sigma^2(\theta)$ jest wariancją asymptotyczną statystyki $\hat{\theta}_n$ i liczba z spełnia

$$\lim_{n\to\infty} P_\theta\left(\left|\frac{\hat{\theta}_n - \theta}{\sigma(\theta)}\sqrt{n}\right| \le z\right) = 0.95.$$

Wyznaczyć postać przedziału ufności.
Odp. $\left[\frac{n\sqrt{n}}{\sum_{i=1}^n X_i^3(\sqrt{n}+1.96)}, \frac{n\sqrt{n}}{\sum_{i=1}^n X_i^3(\sqrt{n}-1.96)}\right]$