Inżynierskie zastosowania statystyki – ćwiczenia

Temat 9: testowanie hipotez statystycznych, estymacja przedziałowa,

Uwaga: Na zajęcia proszę przynieść (i umieć wykorzystać) niezbędne tablice statystyczne rozkładu normalnego, t-Studenta i χ^2.

Zadania z hipotez:

1. W populacji badana cecha ma rozkład normalny N(μ,4). Wylosowano próbkę złożoną z 9 obserwacji z populacji. Zweryfikuj hipotezę $H_0: \mu = 2$ przy hipotezie alternatywnej $H_1: \mu = \mu_1 < 2$ na poziomie istotności α=0,05. Średnia z próbki wynosi $\overline{x} = 1{,}4$.

2 Na dwóch różnych wagach zważono po 10 odcinków przędzy o długości 100 m i uzyskano następujące rezultaty (w g)

a. na wadze nr 1: 5,25; 5,98; 5,83; 5,58; 5,35; 5,59; 5,41; 5,81; 5,95; 5,72
b. na wadze nr 2: 5,31; 5,13; 5,64; 5,89; 5,17; 5,18; 5,27; 5,73; 5,08; 5,24

Wiadomo, że wariancja mas stumetrowych odcinków przędzy dla pierwszej wagi jest równa $\sigma_1^2 = 0{,}06$, a dla drugiej wagi $\sigma_1^2 = 0{,}07$. Zakładając, że rozpatrywana cecha (masa stumetrowego odcinka) ma rozkład normalny na poziomie istotności $1 - \alpha = 0{,}95$, zweryfikuj hipotezę H_0, że wartości przeciętne mas odcinków przędzy uzyskiwane przez te wagi są jednakowe, wobec hipotezy alternatywnej $H_1: \mu_1 \neq \mu_2$.

3. W pewnym województwie przyjmuje się, że plony żyta mają rozkład normalny o nieznanych parametrach. Przyjmuje się, że średni plon z powierzchni wynosi 28 ha/q. Badania przeprowadzono na 20 powierzchniach i otrzymano następujące rezultaty: średni plon 25 q/ha, odchylenie standardowe 4 q/ha. Zweryfikuj na poziomie istotności $\alpha = 0{,}01$ hipotezę zerową $H_0: m = 28$.

4. Pewnym przyrządem pomiarowym wykonano 8 pomiarów wielkości i otrzymano wartości: 18,17; 18,21; 18,05; 18,14; 18,19; 18,22; 18,06; 18,08. Zweryfikuj na poziomie istotności α=0,05 hipotezę $H_0: \sigma^2 = 0{,}06$ dotyczącą wariancji wskazań przyrządu wobec hipotezy alternatywnej $H_1: \sigma^2 \neq 0{,}06$

5. Z populacji, w której badana cecha ma nieznaną dystrybuantę F, pobrano próbkę o liczności 200. Otrzymane wyniki po podziale na 10 równych klas zawarto w dwóch pierwszych kolumnach poniższej tabeli. Na poziomie istotności α=0,05 zweryfikować hipotezę H : {F(x) jest dystrybuantą rozkładu równomiernego na przedziale (45,50)

środki klas	liczności klas n_i
45,25	23
45,75	19
46,25	25
46,75	18
47,25	17
47,75	24
48,25	16
48,75	22
49,25	20
49,75	16

6. W fabryce pracują dwie zmiany robotników. Zmierzono czasu wykonania pewnego elementu na poszczególnych zmianach. Dla zmiany pierwszej otrzymano wartości (w minutach): 12, 13, 18 25, 42, 19, 22, 35. Dla zmiany drugiej: 23, 30, 27, 17, 21, 33, 31. Za pomocą testu serii na poziomie istotności $\alpha = 0{,}05$ zweryfikuj hipotezę, że czasy wykonywania rozpatrywanego detalu na obydwu zmianach mają te same rozkłady przeciw hipotezie alternatywnej, że rozkłady się różnią.

Zadania z przedziałów ufności:

1. Na podstawie n-elementowej próby prostej ($X_1, X_2, \ldots, X_n$) znajdź przedział ufności dla nieznanej wartości przeciętnej μ populacji, w której badana cecha ma rozkład $N(\mu, \sigma)$. Zakładamy, że odchylenie standardowe σ jest znane.

2. W fabryce produkowana jest farba olejna dostarczana w postaci puszek o pojemności 2 litry. Zakładamy, że rozkład ilości farby w puszkach jest normalny z odchyleniem standardowym wynoszącym 11 ml. Dział kontroli jakości przeprowadził losowo 9 pomiarów ilości farby w puszkach napełnianych przez automat i otrzymano następujące wyniki (w ml): 2008, 1990, 1986, 1978, 2002, 1998, 2005, 1975, 1988. Na tej podstawie wyznacz przedział ufności dla nieznanej średniej ilości farby w puszce przyjmując współczynnik ufności równy $1 - \alpha = 0{,}95$.

3. W gospodarstwie na 16 polach doświadczalnych zasadzono nowa odmianę ziemniaków w celu sprawdzenia jej wydajności. Po pierwszych zbiorach otrzymano następujące charakterystyki plonów z ha: $\bar{x} = 264q,\ S = 15q$. Wiadomo, że rozkład plonów z ha w naturalnych warunkach jest w przybliżeniu normalny. Na podstawie uzyskanych wyników wyznacz przedział ufności dla nieznanego średniego plonu z ha nowej odmiany ziemniaków przyjmując współczynnik ufności 0,90.

4. Dyrekcja hipermarketu poleciła oszacować średni czas obsługi klienta przy kasie. Nie jest znana postać rozkładu czasu obsługi oraz nie są znane jego parametry. Wykonano pomiar czasu obsługi klienta dla 200 losowo wybranych osób. W ten sposób otrzymano empiryczne wartości $\bar{x} = 56\ s, S = 8\ s.$ Wyznacz przedział ufności dla średniego czasu obsługi przy współczynniku ufności $1 - \alpha = 0{,}95$ oraz porównaj otrzymaną wartość przy współczynniku ufności $1 - \alpha = 0{,}90$.

5. Dla losowo wybranych odcinków przędzy o długości 1m wykonano pomiary liczby skrętów i uzyskano następujące wyniki: 87, 102, 119, 81, 97, 93, 100, 114, 99, 100, 113, 93, 95, 85, 123, 99. Znajdź 90% realizacje przedziałów ufności wariancji i odchylenia standardowego liczby skrętów całej partii przędzy. Załóż, że liczba skrętów odcinków przędzy ma rozkład normalny.

6. Wyznacz minimalną liczebność próby dla danych z zadania 2 tak, aby maksymalny błąd szacunku średniej przy współczynniku ufności 0,95 wynosił 2 ml.

Przedziały ufności dla średniej

Model I Populacja generalna ma **rozkład *N*(*m*,σ)**, odchylenie standardowe **jest znane**. Nieznany jest parametr m, dla którego szukamy przedziału ufności. Dla próby o liczebności n

$$\Pr\left(\overline{X} - u(1-0.5\alpha)\frac{\sigma}{\sqrt{n}} < m < \overline{X} + u(1-0.5\alpha)\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha.$$

Model II Populacja generalna ma **rozkład *N*(*m*,σ)**, odchylenie standardowe **jest nieznane**. Nieznany jest też parametr m, dla którego szukamy przedziału ufności. Dla próby o liczebności n

$$\Pr\left(\overline{X} - t(1-0.5\alpha, n-1)\frac{S}{\sqrt{n-1}} < m < \overline{X} + t(1-0.5\alpha, n-1)\frac{S}{\sqrt{n-1}}\right) = 1-\alpha.$$

$$\Pr\left(\overline{X} - t(1-0.5\alpha, n-1)\frac{\hat{S}}{\sqrt{n}} < m < \overline{X} + t(1-0.5\alpha, n-1)\frac{\hat{S}}{\sqrt{n}}\right) = 1-\alpha,$$

Model III Rozkład dowolny, ale n musi być duże (co najmniej kilkadziesiąt) oraz istnieje wariancja $\sigma^2 < \infty$, która może być nieznana. Wówczas przedziały ufności, wyznaczane są podobnie jak w modelu I, ale zamiast σ należy podstawić $\hat{S}$.

$$\Pr\left(\overline{X} - u(1-0.5\alpha)\frac{\hat{S}}{\sqrt{n}} < m < \overline{X} + u(1-0.5\alpha)\frac{\hat{S}}{\sqrt{n}}\right) = 1-\alpha.$$

Przedziały ufności dla wariancji

Model I Populacja generalna ma rozkład normalny. Nieznany jest parametr σ, dla którego szukamy przedziału ufności. Próba jest mała (n<50). Przedział ufności określony jest wzorem

$$\Pr\left(\frac{nS^2}{\chi^2(1-0.5\alpha, n-1)} < \sigma^2 < \frac{nS^2}{\chi^2(0.5\alpha, n-1)}\right) = 1-\alpha$$

lub równoważnie

$$\Pr\left(\frac{(n-1)\hat{S}^2}{\chi^2(1-0.5\alpha, n-1)} < \sigma^2 < \frac{(n-1)\hat{S}^2}{\chi^2(0.5\alpha, n-1)}\right) = 1-\alpha$$

Model II

Populacja generalna ma rozkład normalny lub zbliżony do normalnego i próba jest duża (n≥50). Wtedy

$$\Pr\left(\frac{S}{1+\dfrac{u(1-0.5\alpha)}{\sqrt{2n}}} < \sigma^2 < \frac{S}{1+\dfrac{u(1-0.5\alpha)}{\sqrt{2n}}}\right) \approx 1-\alpha.$$