Statystyczna Analiza Danych
SAD-2020-2021

Wykład 8

Estymacja punktowa

Niech $X_1, X_2, ..., X_n$ będzie prostą próbą losową z rozkładu, którego parametr θ jest nieznany.

Definicja

Statystykę $h(X_1, X_2, ..., X_n)$, której realizacje dla konkretnych próbek są „rozsądnymi" ocenami θ, nazywamy **estymatorem** parametru θ i oznaczamy

$$\boxed{\hat{\theta} = h(X_1, X_2, ..., X_n)}$$

Definicja

Estymator $\hat{\theta}$ parametru θ jest nieobciążony, jeśli

$$\boxed{E(\hat{\theta}) = \theta}.$$

Estymacja punktowa

Przykłady

- Średnia z prostej próby losowej jest nieobciążonym estymatorem wartości średniej μ:

$$\boxed{E(\overline{X}) = \mu}$$

- Wariancja z prostej próby losowej jest nieobciążonym estymatorem wariancji rozkładu cechy populacji σ^2:

$$\boxed{E(S^2) = \frac{1}{n-1} E\left(\sum_{i=1}^{n} (X_i - \overline{X})^2 \right) = \sigma^2}$$

Estymacja przedziałowa

$$g, h: (-\infty, \infty) \to (-\infty, \infty), \quad g(\cdot) < h(\cdot)$$

$\alpha \in (0{,}1) - \text{mała liczba, np.}\ \alpha = 0{,}05,\ \ 0{,}01,\ \ 0{,}02,\ \ 0{,}1, \ldots$

Definicja. Przedział losowy

$$[g(X_1, X_2, \ldots, X_n), h(X_1, X_2, \ldots, X_n)]$$

nazywamy przedziałem ufności dla parametru θ na poziomie ufności $1 - \alpha$, jeśli

$$P(\theta \in [g(X_1, X_2, \ldots, X_n), h(X_1, X_2, \ldots, X_n)]) = 1 - \alpha$$

Estymacja przedziałowa

I. Przedziały ufności dla wartości średniej rozkładu normalnego.

Niech $X_1, X_2, ..., X_n$ będzie prostą próbą losową z rozkładu normalnego $N(\mu, \sigma)$.

<u>Model 1</u>. (<u>znane odchylenie standardowe populacji σ</u>)

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \quad \sim \quad N(0,1).$$

Niech $\alpha \in (0,1)$ - ustalona liczba.

Estymacja przedziałowa

$$P(z_{\alpha/2} \le Z \le z_{1-\alpha/2}) = 1-\alpha, \qquad (1)$$

gdzie $z_{\alpha/2}$ = kwantyl rzędu $\alpha/2$ rozkładu $N(0,1)$,

$z_{1-\alpha/2}$ = kwantyl rzędu $1-\alpha/2$ rozkładu $N(0,1)$, tzn.

$$P(Z \le z_{\alpha/2}) = \frac{\alpha}{2} \quad \text{oraz} \quad P(Z \le z_{1-\alpha/2}) = 1-\frac{\alpha}{2}.$$

Z symetrii standardowej gęstości normalnej

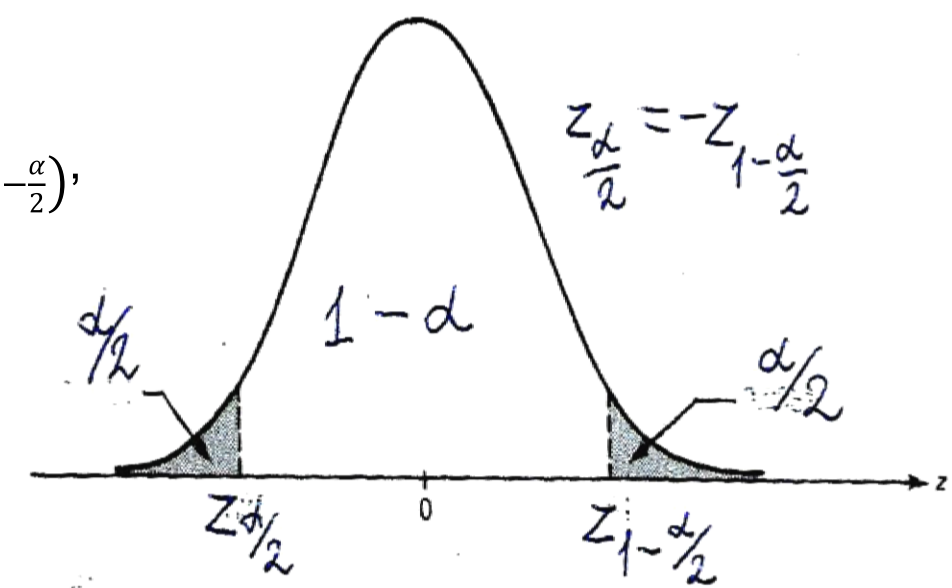
$$z_{\alpha/2} = -z_{1-\alpha/2}.$$

Równanie (1) można zapisać jako

$$P(-z_{1-\alpha/2} \le Z \le z_{1-\alpha/2}) = 1-\alpha.$$

$z_{\frac{\alpha}{2}}$ – kwantyl rzędu $\frac{\alpha}{2}$ rozkładu standardowego normalnego

$$\Phi\left(z_{\frac{\alpha}{2}}\right) = \frac{\alpha}{2} = 1 - \Phi\left(z_{1-\frac{\alpha}{2}}\right),$$



Estymacja przedziałowa

Podstawiając dokładną postać *Z* mamy

$$P(-z_{1-\alpha/2} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le z_{1-\alpha/2}) =$$

$$P(-z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \le \overline{X}-\mu \le z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}) =$$

$$P(\overline{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1-\alpha. \qquad (2)$$

Definicja

Przedział losowy

$$\boxed{[\overline{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]}$$

zawierający z prawdopodobieństwem $1-\alpha$ nieznaną wartość średnią μ = **przedział ufności dla** μ **na poziomie ufności** $1-\alpha$

Estymacja przedziałowa

Definicja

Przedział ufności dla wartości średniej populacji μ na poziomie ufności $1-\alpha$

$$\boxed{[\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]}$$

(obliczony na podstawie konkretnej próbki).

Estymacja przedziałowa

Interpretacja częstościowa (sens praktyczny) przedziału ufności

Niech $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_N$ oznaczają średnie próbkowe obliczone dla N próbek: $(x_1^1, x_2^1, \ldots x_n^1)$, $(x_1^2, x_2^2, \ldots x_n^2)$, … , $(x_1^N, x_2^N, \ldots x_n^N)$.
Próbki są realizacjami niezależnych prostych prób losowych $(X_1^1, X_2^1, \ldots, X_n^1)$, $(X_1^2, X_2^2, \ldots, X_n^2)$,…., $(X_1^N, X_2^N, \ldots, X_n^N)$.
Dokładniej: wykonujemy N jednakowych niezależnych doświadczeń. Każde k-te ($k = 1,2,\ldots,N$) doświadczenie polega na zaobserwowaniu realizacji k-tej prostej próby losowej $(X_1^k, X_2^k, \ldots, X_n^k)$, tzn. k-tej próbki: $(x_1^k, x_2^k, \ldots x_n^k)$. Przedział ufności dla μ na poziomie ufności $1-\alpha$ obliczony dla k-tej próbki ma postać

$$[\bar{x}_k - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x}_k + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}].$$

Estymacja przedziałowa

Nieznana nam średnia μ nie dla każdej próbki należy do wyznaczonego dla niej przedziału ufności.
Ale, niech N_μ oznacza liczbę tych doświadczeń, dla których

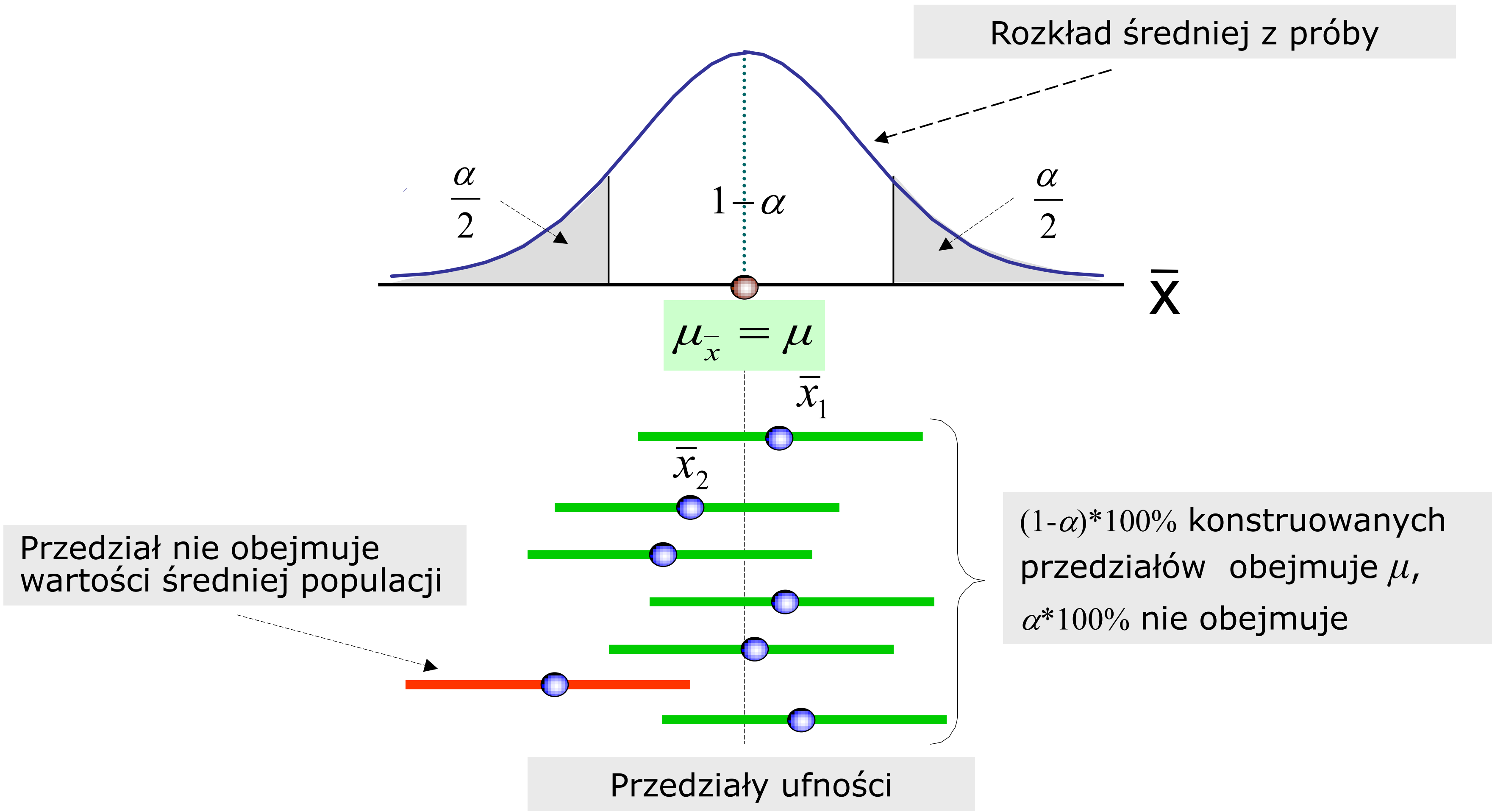
$$\mu \in [\bar{x}_k - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x}_k + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}].$$

Wówczas na mocy interpretacji częstościowej prawdopodobieństwa zdarzenia, dla $N \to \infty$,

$$\frac{N_\mu}{N} \to P(\mu \in [\bar{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]) = 1-\alpha$$

Zatem spośród wielu próbek w przybliżeniu $(1-\alpha)100\%$ jest takich, dla których wyznaczony przedział ufności zawiera nieznaną wartość średnią μ.

Estymacja przedziałowa



Estymacja przedziałowa

Jak duża powinna być liczność próbki n ?

Długość przedziału ufności

$$[\overline{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]$$

jest **stała (nie zależy od próbki)** równa

$$\boxed{2z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}}.$$

Im większe n, tym mniejsza długość przedziału ufności, tzn. tym lepsze (dokładniejsze) oszacowanie przedziałowe μ na danym poziomie ufności. Ze wzoru **(2)** mamy

$$P(\left|\overline{X} - \mu\right| \le z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1-\alpha,$$

Niech $d > 0$ będzie takie że

$$z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \le d, \quad \text{równoważnie} \quad n \ge \left(\frac{z_{1-\alpha/2}\sigma}{d}\right)^2.$$

Estymacja przedziałowa

Wówczas (wykorzystując $P(A) \le P(B)$ dla $A \subset B$)

$$1-\alpha = P(\left|\overline{X}-\mu\right| \le z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}) \le P(\left|\overline{X}-\mu\right| \le d)$$

$$P(\left|\overline{X}-\mu\right| \le d) \ge 1-\alpha$$

Stwierdzenie. Jeśli liczność prostej próby losowej z rozkładu normalnego o wartości średniej μ i standardowym odchyleniu σ spełnia warunek

$$n \ge \left(\frac{z_{1-\alpha/2}\sigma}{d}\right)^2, \qquad \text{to} \qquad P(\left|\overline{X}-\mu\right| \le d) \ge 1-\alpha .$$

(Z prawdopodobieństwem co najmniej $1-\alpha$ błąd bezwzględny oszacowania nieznanej wartości średniej μ poprzez $\overline{X}$ nie przekroczy d, tzn. wśród wielu próbek o liczności *n* częstość takich, dla których błąd bezwzględny średniej próbkowej nie przekroczy *d*, jest w przybliżeniu nie mniejsza niż $1-\alpha$.)

Estymacja przedziałowa

Zadanie

Producent chce ocenić średnią zawartość nikotyny w paczkach papierosów pewnego gatunku. Wiadomo, że standardowe odchylenie zawartości nikotyny w losowo wybranej paczce papierosów $\sigma = 8$ (mg),

Wyznaczyć liczbę paczek papierosów, w których należy zbadać zawartość nikotyny, aby na poziomie ufności co najmniej 0,95 móc stwierdzić, że obliczona średnia z próbki $\overline{x}$ nie będzie się różniła od prawdziwej średniej zawartości nikotyny μ o więcej niż 1,5 (mg).

Estymacja przedziałowa

Zakładając rozkład normalny zawartości nikotyny w paczce papierosów mamy:

$$P(|\overline{X}-\mu| \le d) \ge 0{,}95, \text{ jeśli } n \ge \left(\frac{z_{1-\alpha/2}\sigma}{d}\right)^2,$$

gdzie $\alpha = 0{,}05$, $\sigma = 8$, $d = 1{,}5$, $z_{1-\alpha/2} = z_{0{,}975} = 1{,}96$.

Stąd $n \ge \left(\frac{1{,}96 \times 8}{1{,}5}\right)^2$.

Zatem liczność próbki powinna być: $n \ge 109$.

Estymacja przedziałowa

Zadanie
Stacja paliw sprzedała 8019 litrów gazu w ciągu 9 losowo wybranych dni. Załóżmy, że dzienna ilość sprzedanego gazu ma rozkład normalny o standardowym odchyleniu $\sigma = 90$ (litrów). Skonstruować przedziały ufności dla średniej dziennej sprzedaży gazu na poziomach ufności:
(a) 0,98, **(b)** 0,80.

Mamy: $\sum_{i=1}^{9} x_i = 8019$, $n = 9$, $\sigma = 90$, skąd $\bar{x} = \frac{8019}{9} = 891$.

(a) $\alpha = 0{,}02$, $1 - \alpha/2 = 0{,}99$, $z_{0,99} = 2{,}33$.

98% przedział ufności dla μ:

$$[891 - 2{,}33\frac{90}{\sqrt{9}},\ 891 + 2{,}33\frac{90}{\sqrt{9}}] = [821{,}1\ ;\ 960{,}9]$$

Estymacja przedziałowa

Model 2. (nieznane odchylenie standardowe σ)

W poprzednim modelu wykorzystano statystykę $Z = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}}$

Podstawiając zamiast σ jej estymator $S = \sqrt{S^2}$, gdzie

$S^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$, otrzymujemy zmienną losową

$$\boxed{T = \frac{\overline{X} - \mu}{S / \sqrt{n}}}$$

T ma znany rozkład: ***t*-Studenta** z (n – 1) stopniami swobody.

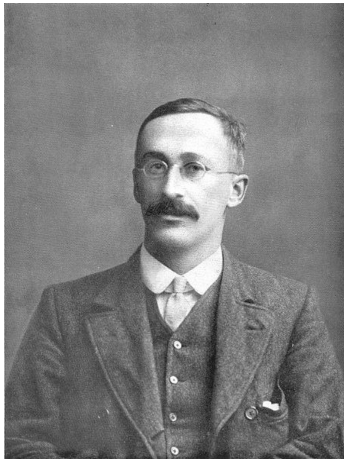
Niech $Z_0, Z_1, \ldots, Z_k$ będą niezależnymi zmiennymi losowymi o rozkładach $N(0,1)$.

Definicja (rozkład *t*-Studenta z *k* stopniami swobody)
Rozkład prawdopodobieństwa zmiennej losowej

$$\boxed{V = \frac{Z_0}{\sqrt{(Z_1^2 + \ldots + Z_k^2)/k}}}$$

nazywamy rozkładem t **Studenta z** k **stopniami swobody**.
Notacja: $V \sim t_k$.

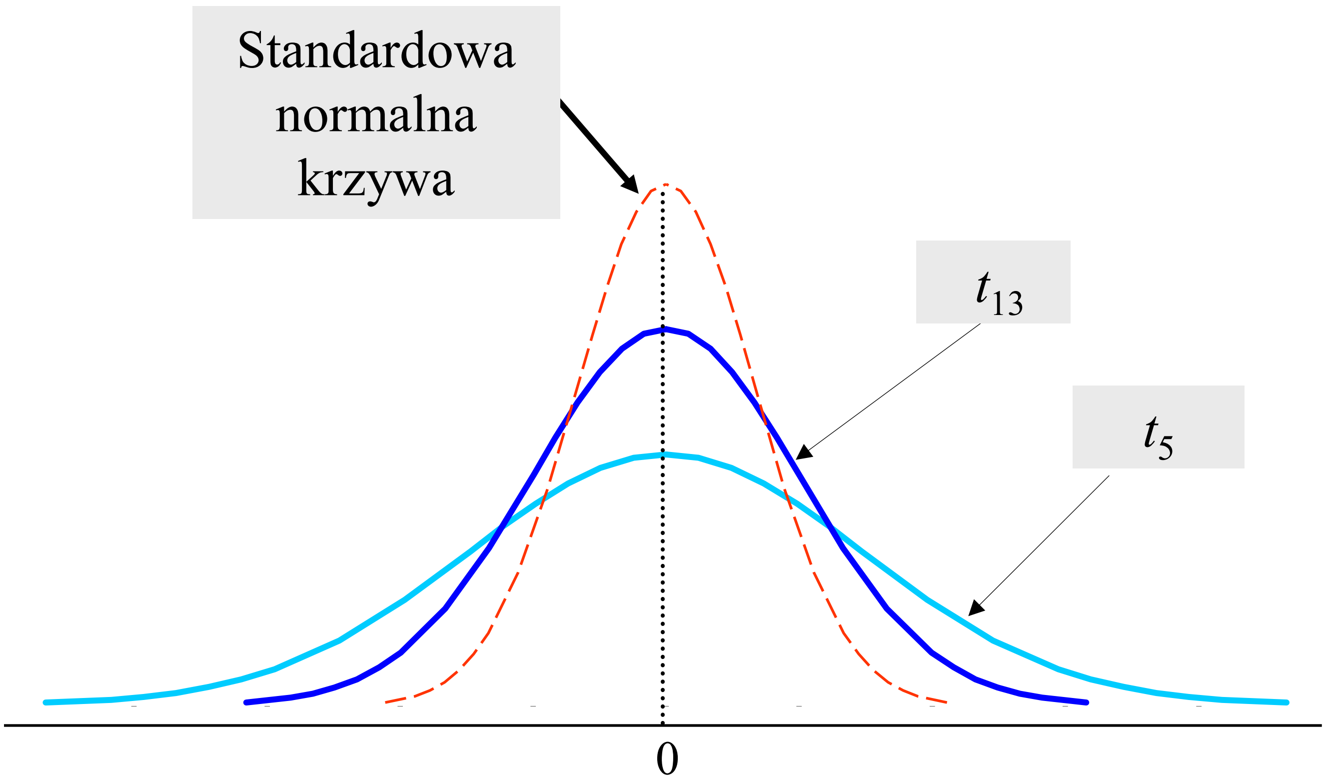
William Sealy Gosset
„Student"
1876 - 1937



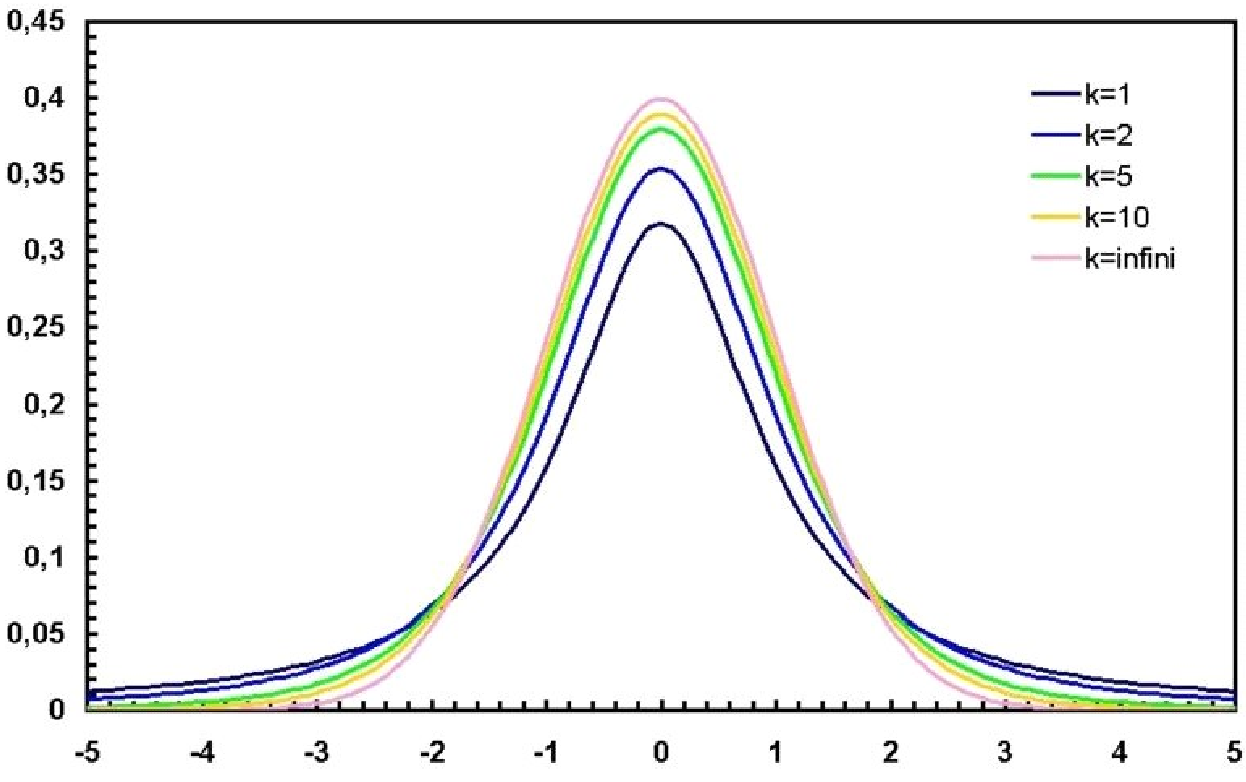
Rozkład t- Studenta

Gęstość symetryczna o podobnym kształcie jak gęstość normalna, $E(V)=0$, Dla $k \geq 30$ można przyjąć

$$V \sim t_k = N(0,1)$$



Gęstości rozkładu t - Studenta



ESTYMACJA PRZEDZIAŁOWA

Mając zmienną losową

$$T = \frac{\overline{X} - \mu}{S / \sqrt{n}} \sim t_{n-1}$$

budujemy przedział ufności dla μ analogicznie jak w modelu 1:

$$\boxed{[\overline{x} - t_{1-\alpha/2,n-1} \frac{s}{\sqrt{n}}, \overline{x} + t_{1-\alpha/2,n-1} \frac{s}{\sqrt{n}}]},$$

gdzie

$t_{1-\alpha/2,n-1}$ = kwantyl rzędu $(1-\alpha/2)$ rozkładu t Studenta o

$(n-1)$ stopniach swobody.

Zadanie. W ciągu pięciu losowo wybranych tygodni zaobserwowano następujące zużycia cukru w gospodarstwie domowym (w kg): 3,8,   4,5,   5,2,   4,0,   5,5. Skonstruować 90% przedział ufności dla średniego tygodniowego zużycia cukru w tym gospodarstwie, jeśli można przyjąć rozkład normalny tygodniowego zużycia cukru.

Obliczamy: $\bar{x} = 4{,}6$ oraz

$$\sum_{i=1}^{5}(x_i - \bar{x})^2 = (-0{,}8)^2 + (-0{,}1)^2 + (0{,}6)^2 + (-0{,}6)^2 + (0{,}9)^2 = 2{,}18$$

$$s^2 = \frac{2{,}18}{5-1} = 0{,}545, \quad s = 0{,}738, \quad \alpha = 0{,}1, \quad 1-\alpha/2 = 0{,}95,$$

liczba stopni swobody: $5 - 1 = 4$, $t_{0,95,4} = 2{,}132.$

Stąd 90% przedział ufności dla μ:

$$[\bar{x} - t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}}, \bar{x} + t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}}]$$

$$= [\ 4{,}6 - 2{,}132\frac{0{,}738}{\sqrt{5}},\ 4{,}6 + 2{,}132\frac{0{,}738}{\sqrt{5}}\ ] = [3{,}896,\ 5{,}304].$$

Przedział ufności dla średniej

Zadanie. Zanotowano czasy obsługi przy okienku kasowym (w minutach) 64 losowo wybranych klientów pewnego banku. Obliczono: średnią z próbki $\bar{x} = 3{,}2$ (min.) oraz wariancję z próbki $s^2 = 1{,}44$ (min.^{2}). Znaleźć 98% przedział ufności dla średniego czasu obsługi μ, jeśli można założyć, że czas obsługi klienta przy okienku kasowym ma rozkład normalny.

duża próba ($n > 30$)

Mamy:

$$\bar{x} = 3{,}2, \quad s = \sqrt{1{,}44} = 1{,}2, \quad n = 64, \quad \alpha = 0{,}02,$$

$$t_{1-0{,}02/2,63} = t_{0{,}99,63} = z_{0{,}99} = 2{,}33.$$

98% przedział ufności dla μ ma postać

$$[\bar{x} - t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}} \quad ; \quad \bar{x} + t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}}]$$

$$[3{,}2 - 2{,}33\frac{1{,}2}{\sqrt{64}} \quad ; \quad 3{,}2 + 2{,}33\frac{1{,}2}{\sqrt{64}}] = [2{,}85 ; 3{,}6]$$

Przedział ufności dla wariancji

Model 3. Przedział ufności dla wariancji.

Niech $X_1, X_2, \ldots, X_n$ będzie prostą próbą losową z rozkładu <u>normalnego</u> $N(\mu, \sigma)$.

Definicja

Niech $Z_1, Z_2, \ldots, Z_n$ będą niezależnymi zmiennymi losowymi o rozkładach $N(0,1)$. Wówczas zmienna losowa

$$\boxed{\chi^2 = \sum_{i=1}^{n} Z_i^2}$$

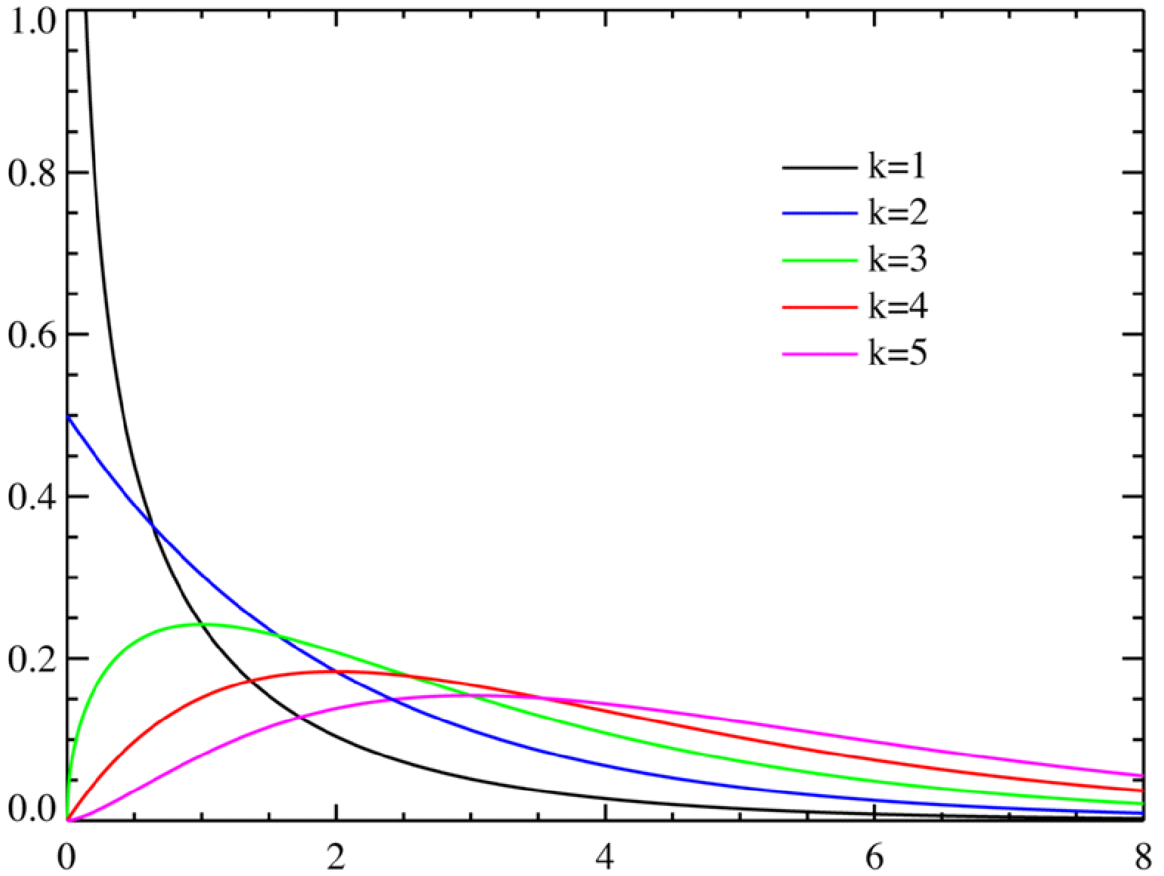
ma rozkład chi – kwadrat o n stopniach swobody, co zapisujemy:

$$\chi^2 \sim \chi_n^2$$

W modelu: $\frac{X_1 - \mu}{\sigma}, \frac{X_2 - \mu}{\sigma}, \ldots, \frac{X_n - \mu}{\sigma}$ są niezależne i mają rozkłady *N*(0,1), zatem

$$\boxed{\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi_n^2}$$

Rozkład chi – kwadrat



Funkcja gęstości prawdopodobieństwa rozkładu chi – kwadrat
o różnej liczbie stopni swobody

Przedział ufności dla wariancji

Zastępując nieznaną wartość średnią μ przez średnią z próby losowej $\overline{X}$ otrzymamy zmienną losową:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{X_i - \overline{X}}{\sigma} \right)^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1} \quad .$$

Stąd

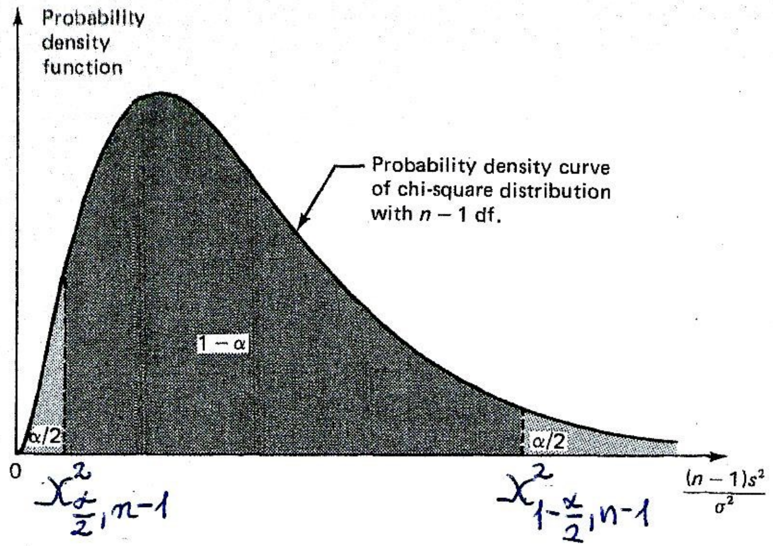
$$P\left(\chi^2_{\alpha/2,n-1} \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{1-\alpha/2,n-1} \right) = 1-\alpha, \qquad \textbf{(3)}$$

gdzie $\chi^2_{\alpha/2,n-1}$, $\chi^2_{1-\alpha/2,n-1}$ są odpowiednio kwantylami rzędu $\alpha/2$, i $(1-\alpha/2)$, rozkładu χ^2_{n-1}.

Wzór **(3)** zapisujemy równoważnie:

$$P\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}} \right) = 1-\alpha.$$

Rozkład chi-kwadrat



Chi-square values such that area $\alpha/2$ is to the right and area $\alpha/2$ is to the left.

Przedział ufności dla wariancji

Stąd, przedziałami ufności na poziomie ufności $(1-\alpha)$ są:

- dla wariancji σ^2 rozkładu normalnego

$$\boxed{\left[\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}, \frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}\right]}$$

- dla standardowego odchylenia σ rozkładu normalnego

$$\boxed{\left[\sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}}, \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}}\right]}$$

III. Przedziały ufności dla proporcji

Model 4. Niech $X_1, X_2, \ldots, X_n$ będzie prostą próbą losową z rozkładu Bernoulli'ego o nieznanym parametrze p.

Wówczas $\mu = E(X_i) = p, \quad \sigma^2 = p(1-p)$.

Niech $\hat{p} = \overline{X}$. Z centralnego twierdzenia granicznego dla dostatecznie dużego *n* zmienna losowa $\dfrac{\hat{p}-p}{\sqrt{\dfrac{p(1-p)}{n}}}$ ma rozkład bliski rozkładowi $N(0,1)$. (musi zachodzić $np \geq 5, n(1-p) \geq 5$)

Podobnie $\dfrac{\hat{p}-p}{\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}}$ ma rozkład bliski $N(0,1)$ o ile $n\hat{p} \geq 5, n(1-\hat{p}) \geq 5$.

Przedział ufności dla proporcji

Stąd

$$P\left(-z_{1-\alpha/2} \le \frac{\hat{p}-p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \le z_{1-\alpha/2}\right) \approx 1-\alpha$$

Równoważnie

$$P\left(\hat{p}-z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p}+z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = \approx 1-\alpha$$

Przedział ufności dla proporcji p na poziomie ufności $(1-\alpha)$ jest postaci:

$$\boxed{\left[\hat{p}-z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\quad \hat{p}+z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]}$$

Przedział ufności dla proporcji

Przykład. W badaniach opinii publicznej otrzymano wynik: 70% spośród 1000 ankietowanych Polaków uważa, że wejście Polski do Unii Europejskiej jest dla nas korzystne, a pozostałych 30% osób sądzi, że nie bądź nie ma zdania. Skonstruować 95% przedział ufności dla proporcji *p* Polaków, którzy uważają, że obecność Polski w UE jest korzystna.

Mamy: $\hat{p} = 0,7, \quad 1-\alpha = 0,95, \quad \alpha = 0,05, \quad 1-\alpha/2 = 1-0,025 = 0,975$.

Z tablicy dystrybuanty rozkładu normalnego: $z_{0,975} = 1,96$,

$n\hat{p} = 1000 \times 0,7 \geq 5 \quad n(1-\hat{p}) = 1000 \times 0,3 \geq 5$.

Zatem można wykorzystać przybliżony przedział ufności:

$$\left[\hat{p} - z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\ ;\ \hat{p} + z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right] =$$

$$\left[0,7 - 1,96\sqrt{\frac{0,7\times 0,3}{1000}}\ ;\ 0,7 + 1,96\sqrt{\frac{0,7\times 0,3}{1000}}\right] = [0,68 \quad ; \quad 0,73].$$

Zatem mamy „95% pewności (ufności)”, że proporcja Polaków uważających za korzystne wejście Polski do UE jest liczbą z przedziału $[0,68 ; 0,73]$.

ESTYMACJA PRZEDZIAŁOWA

Zadanie. Spośród 400 dorosłych przypadkowo wybranych osób zapytanych o regularne uprawianie sportu rekreacyjnego 160 osób odpowiedziało twierdząco. Skonstruować 98% przedział ufności dla proporcji osób uprawiających sport rekreacyjny w danej populacji (*p*).

Mamy:

$$\hat{p} = \frac{160}{400} = 0{,}4, \quad n = 400, \quad \alpha = 0{,}02, \quad 1 - \alpha/2 = 0{,}99, \quad z_{0,99} = 2{,}33.$$

$$n\hat{p} = 160 \geq 5, \quad n(1-\hat{p}) = 240 \geq 5.$$

98% przedział ufności dla proporcji *p*.

$$\left[0{,}4 - 2{,}33\sqrt{\frac{0{,}4 \times 0{,}6}{400}} \;;\; 0{,}4 + 2{,}33\sqrt{\frac{0{,}4 \times 0{,}6}{400}}\right] = [0{,}343\,;\; 0{,}457]$$

Estymacja przedziałowa

Zadanie. Producent twierdzi, że niezawodność elementów jego produkcji wynosi 0,9 (np. prawdopodobieństwo poprawnej pracy w okresie gwarancji, prawdopodobieństwo wylosowania elementu nie spełniającego norm z bieżącej produkcji ..., ogólnie proporcja elementów niezawodnych). Wśród 100 wybranych losowo elementów 15 okazało się zawodnych. Czy jesteśmy skłonni ufać twierdzeniu producenta ?

Estymacja przedziałowa

Przedział ufności dla proporcji *p* na poziomie ufności 0,95:

$$[0{,}85 - 1{,}96 \times \sqrt{\frac{0{,}85 \times 0{,}15}{100}}, 0.85 + 1{,}96 \times \sqrt{\frac{0{,}85 \times 0{,}15}{100}}]$$

= [0,780,0,920] = 95% przedział ufności dla proporcji elementów niezawodnych.

Zasady etyczne podawania analiz statystycznych

- Przedział ufności należy wyznaczyć oprócz estymatora punktowego
- Poziom ufności musi być podany
- Liczność próby musi być podana
- Interpretacja przedziału ufności koniecznie podana