4. TESTOWANIE HIPOTEZ PARAMETRYCZNYCH

Teoria weryfikacji hipotez statystycznych – zajmuje się tworzeniem reguł umożliwiających rozstrzyganie o słuszności sądów (hipotez statystycznych).

Hipotezy statystyczne są sądami dotyczącymi rozkładu populacji; *hipotezy parametryczne* dotyczą wartości parametrów rozkładu – są najczęściej sprawdzanymi hipotezami statystycznymi, *hipotezy nieparametryczne* są np. opiniami dotyczącymi typu rozkładu czy przypuszczeniami o jednakowym rozkładzie dwóch populacji.

Proces weryfikacji hipotezy polega na:

1. postawieniu *hipotezy zerowej* H_0 (poddawanej weryfikacji) i *hipotezy alternatywnej* H_1 (przeciwstawnej weryfikowanej hipotezie H_0),
2. wyznaczeniu statystyki testowej (zależnej od postawionych hipotez),
3. wyznaczeniu w oparciu o przyjęty *poziom istotności* α tzw. *obszaru krytycznego* (jest to obszar odrzucenia hipotezy H_0), *lub*

   wyznaczeniu *granicznego poziomu istotności* (ang. *p–value*, pol. *p–wartość*) tzn. najmniejszego *poziomu istotności* przy którym hipoteza H_0 może zostać odrzucona,
4. podjęcie decyzji o odrzuceniu hipotezy H_0 w sytuacji gdy wartość statystyki testowej należy do *obszaru krytycznego* lub

   gdy *poziom istotności* α jest większy od *poziomu granicznego p–value*

   (jeżeli wartość statystyki nie należy do obszaru krytycznego to nie ma podstaw do odrzucenia hipotezy lub gdy *poziom istotności* α jest mniejszy od *granicznego poziomu istotności p–value*).

W przypadku *hipotez parametrycznych*, sprawdzających wartość nieznanego parametru rozkładu:

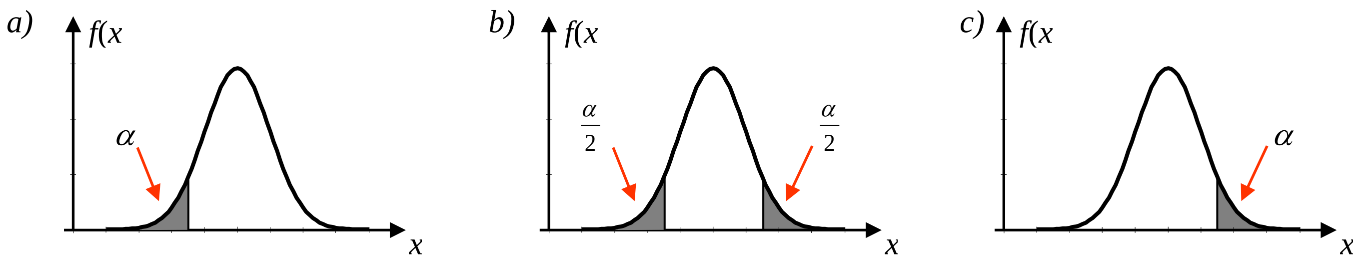
- *hipoteza zerowa* H_0 zakłada, że wartość nieznanego parametru rozkładu θ wynosi θ_0, czyli:

  H_0: $\theta = \theta_0$,
- *hipoteza alternatywna* $H1$ ma najczęściej jedną z postaci:

  a) H_1: $\theta < \theta_0$,   *b)* H_1: $\theta \neq \theta_0$,   *c)* H_1: $\theta > \theta 0$,
- *statystyka testowa* jest estymatorem weryfikowanego parametru rozkładu,
- *obszar krytyczny* jest wyznaczany w oparciu o rozkład przyjętej statystyki testowej, założony *poziom istotności* α i postać *hipotezy alternatywnej,*

  hipoteza zerowa jest odrzucana jeżeli wartość statystyki testowej przy założeniu prawdziwości *hipotezy zerowej* (tzn. wartość estymatora weryfikowanego parametru rozkładu) należy do *obszaru krytycznego,*
- *graniczny poziom istotności p–value* jest wyznaczany w oparciu o rozkład przyjętej statystyki testowej i jej wartość przy założeniu prawdziwości *hipotezy zerowej,*

  hipoteza zerowa jest odrzucana jeżeli *poziom istotności* α jest większy od *granicznego poziomu istotności p–value.*



Rys. 1. Obszar krytyczny a) lewostronny H_1: $\theta < \theta_0$, b) obustronny H_1: $\theta \neq \theta_0$, c) prawostronny H_1: $\theta > \theta_0$.

4.1. Weryfikacja hipotez dla średniej

Statystyką testową wykorzystywaną podczas weryfikacji hipotez o wartości średniej w populacji μ jest średnia z próby x.

Rozkład statystyki testowej, czyli w tym przypadku średniej z próby, jest zależny od posiadanej wiedzy o populacji:

- jeżeli populacja generalna ma rozkład $\mathcal{N}(\mu, \sigma)$ przy czym <u>znane jest</u> odchylenie standardowe populacji σ to rozkład średniej z próby jest zbieżny do rozkładu $\mathcal{N}(\mu, \sigma/\sqrt{n})$, do weryfikacji hipotezy o średniej wykorzystywana jest unormowana zmienna o rozkładzie $\mathcal{N}(0, 1)$:

$$U = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{\bar{x} - \mu}{\sigma}\sqrt{n},$$

- jeżeli populacja generalna ma rozkład $\mathcal{N}(\mu, \sigma)$ przy czym <u>nie jest znane</u> odchylenie standardowe populacji σ (i nie można przyjąć, że $\sigma \approx s$ ponieważ n jest małe) to po podstawieniu odchylenia z próby do statystyki testowej otrzymuje się jej nową postać (patrz: przedział ufności dla μ):

$$t = \frac{\bar{x} - \mu}{s}\sqrt{n},$$

  statystyka ta ma rozkład t – *Studenta* o $(n - 1)$ stopniach swobody.

Przykład 1.

Rozkład pomiarów długości detalu jest rozkładem normalnym z odchyleniem standardowym $\sigma = 1.5$. Na podstawie 10 niezależnych pomiarów długości detalu wyznaczono jego średnią długość równą ok. 20.34. Zweryfikować na poziomie istotności $\alpha = 0.01$ hipotezę, że rzeczywista długość detalu wynosi 20.

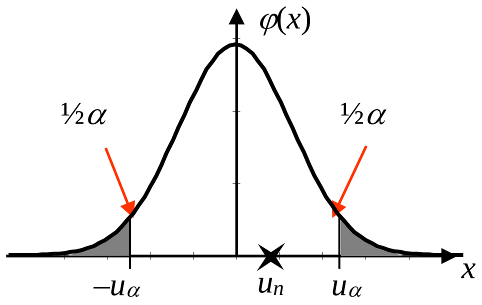
1. Należy zweryfikować *hipotezę zerową* H_0: $\mu = 20$, wobec *hipotezy alternatywnej* H_1: $\mu \neq 20$.
2. W przypadku gdy znane jest odchylenie standardowe do zweryfikowania hipotezy przyjmowana jest statystyka testowa:

$$U = \frac{\bar{x} - \mu}{\sigma}\sqrt{n}.$$

Obliczona na podstawie wyników z próby wartość statystyki testowej (przy założeniu, że hipoteza H_0 jest prawdziwa) wynosi:

$$u_n = \frac{20.34 - 20}{1.5}\sqrt{10} \approx 0.72.$$

3. *Obszar krytyczny* wyznacza się korzystając z odwrotności dystrybuanty rozkładu $\mathcal{N}(0,1)$:

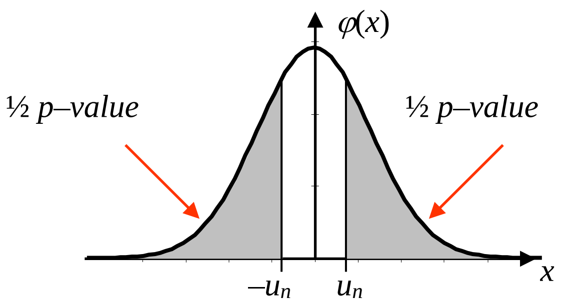


$$u_\alpha = -\Phi^{-1}\left(\frac{\alpha}{2}\right) = -\Phi^{-1}\left(\frac{0.01}{2}\right) \approx 2.58.$$

4. Ponieważ obliczona wartość statystyki testowej znajduje poza *obszarem krytycznym* ($|u_n| < |u_\alpha|$) więc nie ma podstaw do odrzucenia hipotezy H_0.

Identyczny rezultat otrzymuje się wyznaczając *graniczny poziom istotności p–value*.

3. *Graniczny poziom istotności p–value* wyznacza się korzystając z dystrybuanty rozkładu $\mathcal{N}(0,1)$:

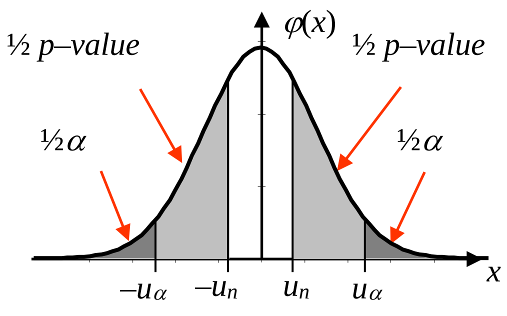


$$\tfrac{1}{2}\, p\text{–}value = \Phi(-u_n) = \Phi(-0.72) \approx 0.24$$

$$\Downarrow$$

$$p\text{–}value \approx 0.48$$

4. Ponieważ założony *poziom istotności* α jest mniejszy od obliczonej wartości *p–value* więc nie ma podstaw do odrzucenia hipotezy H_0.



Przykład 2.

Automat produkuje detale o nominalnej długości 20. Wykonano 10 niezależnych pomiarów długości pewnego detalu otrzymując wyniki: 20, 21, 22.4, 23, 21.3, 21.9, 20.6, 21, 19.8, 20.4. Czy obliczona średnia długość detalu równa 21.14 pozwala na stwierdzenie, że rzeczywista długość jest większa od nominalnej. Przyjąć poziom istotności $\alpha = 0.01$.

1. Należy zweryfikować *hipotezę zerową* H_0: $\mu = 20$, wobec *hipotezy alternatywnej* H_1: $\mu > 20$.

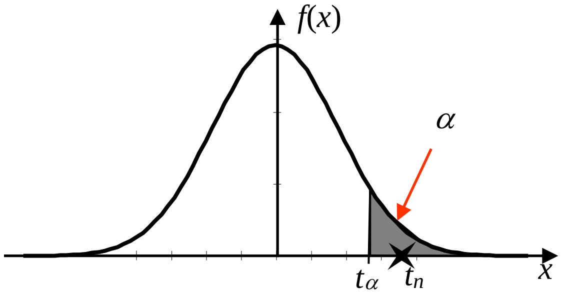
2. W przypadku gdy odchylenie standardowe populacji nie jest znane, do zweryfikowania hipotezy przyjmowana jest statystyka testowa:

$$t = \frac{\bar{x} - \mu}{s}\sqrt{n},$$

   Obliczona na podstawie wyników z próby wartość statystyki testowej (przy założeniu, że hipoteza H_0 jest prawdziwa) wynosi:

$$t_n = \frac{21.14 - 20}{1.03}\sqrt{10} \approx 3.49.$$

3. *Obszar krytyczny* wyznacza się korzystając z odwrotności dystrybuanty rozkładu *t – Studenta* o $(n-1)$ stopniach swobody:



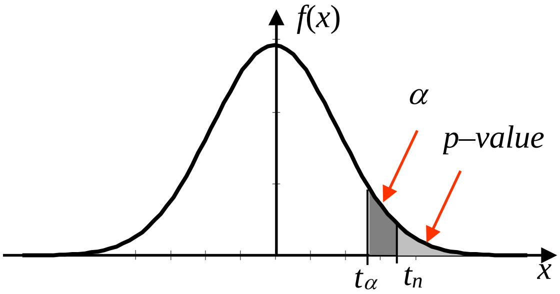
$$t_\alpha = F^{-1}_{t(9)}(1-\alpha) = F^{-1}_{t(9)}(0.99) \approx 2.82.$$

$$t_\alpha = -F^{-1}_{t(9)}(\alpha) = -F^{-1}_{t(9)}(0.01) \approx 2.82.$$

4. Ponieważ obliczona wartość statystyki testowej znajduje wewnątrz *obszaru krytycznego* ($t_n > t_\alpha$) więc hipotezę H_0 należy odrzucić na korzyść hipotezy alternatywnej H_1 – co oznacza, że z prawdopodobieństwem błędu mniejszym od 0.01 można twierdzić, że długość detalu jest większa od nominalnej.

Identyczny rezultat otrzymuje się wyznaczając *graniczny poziom istotności p–value*.

3. *Graniczny poziom istotności p–value* wyznacza się korzystając z dystrybuanty rozkładu *t–Studenta* o $(n–1)$ stopniach swobody:



$$p\text{–}value = F_{t(n-1)}(-t_n) = F_{t(9)}(-3.49) \approx 0.003$$

4. Ponieważ założony *poziom istotności* α jest większy od obliczonej wartość *p–value* więc hipotezę H_0 należy odrzucić.

4.2. Weryfikacja hipotez o równości średnich dwóch populacji

Hipotezy dotyczące równości dwóch średnich sprawdzane są w celu wyjaśnienia przypadkowej lub nieprzypadkowej rozbieżności pewnego parametru populacji. W takim przypadku wykonywane są dwie serie pomiarów, dla każdej z nich obliczana jest średnia. Weryfikacja hipotezy sprowadza się do zbadania różnicy pomiędzy wyznaczonymi średnimi.

Rozkład statystyki testowej, jest zależny od posiadanej wiedzy o populacjach:

- jeżeli populacje generalne mają rozkład odpowiednio: $\mathcal{N}(\mu_1, \sigma_1)$ i $\mathcal{N}(\mu_2, \sigma_2)$ przy czym znane są odchylenia standardowe populacji σ_1 i σ_2, do weryfikacji hipotezy o równości średnich wykorzystywana jest zmienna o rozkładzie $\mathcal{N}(0, 1)$:

$$U = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$$

- jeżeli populacje generalne mają rozkład odpowiednio: $\mathcal{N}(\mu_1, \sigma_1)$ i $\mathcal{N}(\mu_2, \sigma_2)$ przy czym odchylenia standardowe populacji σ_1 i σ_2 nie są znane ale są równe ($\sigma_1 = \sigma_2$), do weryfikacji hipotezy o równości średnich wykorzystywana jest zmienna o rozkład *t – Studenta* o $(n_1 - n_2 - 2)$ stopniach swobody:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \qquad \text{gdzie: } s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}}.$$

Przykład 3.

Wykonano dwie serie pomiarów długości detalu z jednakową dokładnością $\sigma_1 = \sigma_2 = 1.5$. W pierwszej serii przeprowadzono 10 pomiarów: 18, 21, 22.4, 23, 21.3, 21.9, 17.6, 21, 17.8, 19.4 a w drugiej 8 pomiarów: 22.1, 20.3, 21.4, 23.1, 21.1, 21.8, 20.6, 22.8. Zweryfikować na poziomie istotności $\alpha = 0.01$ hipotezę, że rozbieżność średnich jest nieprzypadkowa.

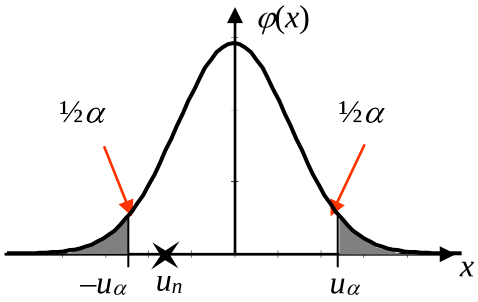
1. Należy zweryfikować *hipotezę zerową* H_0: $\mu_1 = \mu_2$, wobec *hipotezy alternatywnej* H_1: $\mu_1 \neq \mu_2$.
2. W przypadku gdy znane są odchylenia standardowe do zweryfikowania hipotezy przyjmowana jest statystyka testowa:

$$U = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$$

Obliczona na podstawie wyników z próby wartość statystyki testowej (przy założeniu, że hipoteza H_0 jest prawdziwa) wynosi ($\bar{x}_1 = 20.34$, $\bar{x}_2 = 21.65$):

$$u_n = \frac{20.34 - 21.65}{\sqrt{\frac{1.5^2}{10} + \frac{1.5^2}{8}}} = \frac{-1.31}{1.5\sqrt{\frac{1}{10} + \frac{1}{8}}} \approx -1.84.$$

3. *Obszar krytyczny* wyznacza się korzystając z odwrotności dystrybuanty rozkładu $\mathcal{N}(0,1)$:

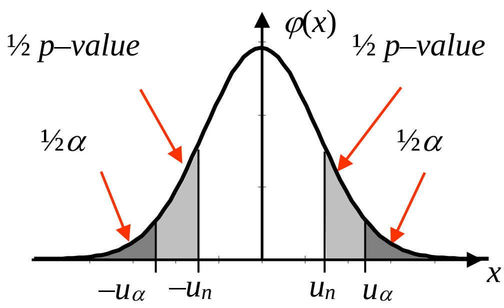


$$u_\alpha = -\Phi^{-1}\left(\frac{\alpha}{2}\right) = -\Phi^{-1}\left(\frac{0.01}{2}\right) \approx 2.58.$$

4. Ponieważ obliczona wartość statystyki testowej znajduje poza *obszarem krytycznym* ($|u_n| < |u_\alpha|$) więc nie ma podstaw do odrzucenia hipotezy H_0.

Analogiczny wniosek otrzymuje się wyznaczając *graniczny poziom istotności p–value*.

3. *Graniczny poziom istotności p–value* wyznacza się korzystając z dystrybuanty rozkładu $\mathcal{N}(0,1)$:



$$½\ p\text{–}value = \Phi(u_n) = \Phi(-1.84) \approx 0.033$$

$$\Downarrow$$

$$p\text{–}value \approx 0.066$$

4. Ponieważ założony *poziom istotności* α jest mniejszy od obliczonej wartości *p–value* więc nie ma podstaw do odrzucenia hipotezy H_0.

Przykład 4.

Rozwiązać zadanie z przykładu 3. zakładając, że odchylenia standardowe populacji σ_1 i σ_2 nie są znane ale są równe ($\sigma_1 = \sigma_2$).

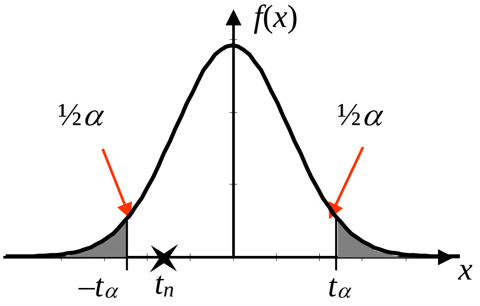
1. Należy zweryfikować *hipotezę zerową* H_0: $\mu_1 = \mu_2$, wobec *hipotezy alternatywnej* H_1: $\mu_1 \neq \mu_2$.
2. W przypadku gdy nie są znane odchylenia standardowe do zweryfikowania hipotezy przyjmowana jest statystyka testowa:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}, \qquad \text{gdzie: } s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}}.$$

Obliczona na podstawie wyników z próby wartość statystyki testowej (przy założeniu, że hipoteza H_0 jest prawdziwa) wynosi:

$$\bar{x}_1 = 20.34,\ \bar{x}_2 = 21.65,\ s_1 \approx 2.00,\ s_2 \approx 1.00,\ s \approx 1.64,\ t_n = \frac{20.34 - 21.65}{1.64\sqrt{\dfrac{1}{10} + \dfrac{1}{8}}} \approx -1.69.$$

3. *Obszar krytyczny* wyznacza się korzystając z odwrotności dystrybuanty rozkładu *t–Studenta* o (n_1+n_2-2) stopniach swobody

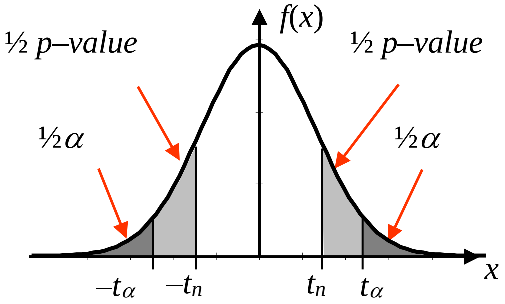


$$t_\alpha = -F^{-1}_{t(16)}\left(\frac{\alpha}{2}\right) = F^{-1}_{t(16)}\left(\frac{0.01}{2}\right) \approx 2.92.$$

4. Ponieważ obliczona wartość statystyki testowej znajduje poza *obszarem krytycznym* ($|t_n| < |t_\alpha|$) więc nie ma podstaw do odrzucenia hipotezy H_0.

Analogiczny wniosek otrzymuje się wyznaczając *graniczny poziom istotności p–value*.

3. *Graniczny poziom istotności p–value* wyznacza się korzystając z dystrybuanty *t–Studenta* o (n_1+n_2-2) stopniach swobody



$$½\ p\text{–}value = F_{t(16)}(-t_n) = F_{t(16)}(-1.69) \approx 0.055$$

$$\Downarrow$$

$$p\text{–}value \approx 0.11$$

4. Ponieważ założony *poziom istotności* α jest mniejszy od obliczonej wartości *p–value* więc nie ma podstaw do odrzucenia hipotezy H_0.

4.3. Weryfikacja hipotez o równości średnich – test dla prób zależnych

W przypadku gdy serie pomiarów wykonywane są dwukrotnie w różnych warunkach, np. przed i po wprowadzeniu pewnej modyfikacji. Weryfikacja hipotezy, tak jak w punkcie poprzednim, sprowadza się do zbadania różnicy pomiędzy wyznaczonymi średnimi – test ten jest właściwie testem dla jednej średniej dla zmiennej D reprezentującej różnicę pomiędzy pomiarami wykonanymi w pierwszych i drugich warunkach:

$$D = X_1 - X_2.$$

Test nazywany jest testem dla populacji zależnych lub sparowanych w odróżnieniu od testów omówionych w punkcie poprzednim dla populacji niezależnych (niesparowanych). Podobnie jak w przypadku weryfikacji hipotezy dla jednej średniej zakłada się, że zmienna D reprezentująca różnice pomiędzy obserwacjami ma rozkład $\mathcal{N}(\mu_d, \sigma_d)$, średnia z różnic $\overline{d}$ ma więc rozkład $\mathcal{N}(\mu_d, \sigma_d/\sqrt{n})$. Do testowania hipotezy wykorzystywane są zmienne:

- jeżeli znane jest odchylenie standardowe σ_d wykorzystywana jest unormowana zmienna o rozkładzie $\mathcal{N}(0, 1)$:

$$U = \frac{\bar{d} - \mu_d}{\sigma_d}\sqrt{n},$$

- jeżeli nie jest znane odchylenie standardowe σ_d (i nie można przyjąć, że $\sigma_d \approx s_d$ ponieważ n jest małe) to po podstawieniu odchylenia z próby do statystyki testowej otrzymuje się jej nową postać

$$t = \frac{\bar{d} - \mu_d}{s_d}\sqrt{n}$$

statystyka ta ma rozkład *t – Studenta* o $(n – 1)$ stopniach swobody,

gdzie: $\bar{d} = \frac{1}{n} d_i$, d_i – różnica pomiędzy obserwacjami dla i–tej pary, łatwo można pokazać, że: $\bar{d} = \bar{x}_1 - \bar{x}_2$, s_d – odchylenie standardowe różnic, n – liczebność próby.

Przykład 5.

Wykonano dwie serie pomiarów twardości 10 losowo wybranych detali. Pomiary wykonywano z wykorzystaniem dwóch różnych wgłębników. Zakładając, że odchylenie standardowe różnicy pomiarów nie jest znane, zweryfikować na poziomie istotności α = 0.01 hipotezę, że rozbieżność średnich jest nieprzypadkowa.

detal	1	2	3	4	5	6	7	8	9	10
wgłębnik1	127	116	128	120	134	104	125	114	123	119
wgłębnik2	111	101	95	94	100	95	106	99	100	105

1. Należy zweryfikować *hipotezę zerową* H_0: $\mu_1 = \mu_2$, wobec *hipotezy alternatywnej* H_1: $\mu_1 \neq \mu_2$.
2. Do zweryfikowania hipotezy przyjmowana jest statystyka testowa:

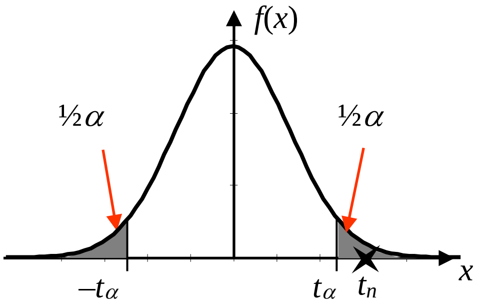
$$t = \frac{\bar{d} - \mu_d}{s_d}\sqrt{n},$$

Po wyznaczeniu różnic dla każdej pary obserwacji:

detal	1	2	3	4	5	6	7	8	9	10
d_i	16	15	33	26	34	9	19	15	23	14

obliczona na podstawie wyników z próby wartość statystyki testowej (przy założeniu, że hipoteza H_0 jest prawdziwa, tzn. średnie są równe czyli spodziewana różnica średnich μ_d wynosi 0) wynosi ($\bar{d} = 20.4$, $s_d = 8.3825$):

$$t_n = \frac{20.4}{8.3825}\sqrt{10} \approx 7.7.$$



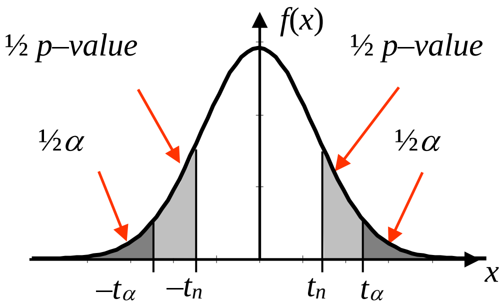
$$t_\alpha = -F^{-1}_{t(9)}\left(\frac{\alpha}{2}\right) = F^{-1}_{t(9)}\left(\frac{0.01}{2}\right) \approx 3.25.$$

3. *Obszar krytyczny* wyznacza się korzystając z odwrotności dystrybuanty rozkładu *t – Studenta* o $(n-1)$ stopniach swobody:

4. Ponieważ obliczona wartość statystyki testowej znajduje wewnątrz *obszaru krytycznego* ($|t_n| > |t_\alpha|$) więc hipotezę H_0 należy odrzucić na korzyść hipotezy alternatywnej H_1 – co oznacza, że z prawdopodobieństwem błędu mniejszym od 0.01 można twierdzić, że pomiary wykonane przy pomocy obydwu wgłębników różnią się od siebie w sposób istotny.

Analogiczny wniosek otrzymuje się wyznaczając *graniczny poziom istotności p–value*.

3. *Graniczny poziom istotności p–value* wyznacza się korzystając z dystrybuanty *t–Studenta* o $(n–1)$ stopniach swobody



$$\tfrac{1}{2}\ p\text{–}value = F_{t(9)}(-t_n) = F_{t(9)}(-7.7) \approx 1.5\mathrm{e}^{-5}$$

$$\Downarrow$$

$$p\text{–}value \approx 0.00003$$

4. Ponieważ założony *poziom istotności* α jest większy od obliczonej wartości *p–value* więc hipotezę H_0 należy odrzucić.

4.4. Weryfikacja hipotez o wariancji

Jeżeli populacja generalna ma rozkład $\mathcal{N}(\mu,\ \sigma)$ przy czym nieznane są średnia i wariancja populacji, wariancję można szacować na postawie próby wykorzystując zależność:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

Do weryfikacji hipotezy o wariancji wykorzystywana jest zmienna o rozkładzie χ^2 o $v = n - 1$ stopniach swobody (n – liczebność wylosowanej próby):

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}.$$

Rozkład zmiennej można pokazać podstawiając do powyższego wzoru zależność opisującą wariancję:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{n-1}{\sigma^2}\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\sigma}\right)^2 = \sum_{i=1}^{n}\left(\frac{(X_i-\mu)-(\overline{X}-\mu)}{\sigma}\right)^2 = \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 +$$

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

$$= \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 - 2n\left(\frac{\overline{X}-\mu}{\sigma}\right)^2 + n\left(\frac{\overline{X}-\mu}{\sigma}\right)^2 = \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 - \left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)^2.$$

Suma $\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2$ jest sumą kwadratów n zmiennych o rozkładzie $\mathcal{N}(0,\ 1)$ więc zgodnie z definicją ma rozkład χ^2 o $v = n$ stopniach swobody, $\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)^2$ jest kwadratem zmiennej o rozkładzie $\mathcal{N}(0,\ 1)$ więc ma rozkład χ^2 o $v = 1$ stopniu swobody. Ostatecznie zmienna $\chi^2 = (n-1)s^2/\sigma^2$ ma rozkład χ^2 o $v = n - 1$ stopniach swobody. ■

4.5. Weryfikacja hipotez o równości wariancji dwóch populacji

Hipotezy dotyczące równości dwóch wariancji sprawdzane są w celu wyjaśnienia przypadkowej lub nieprzypadkowej rozbieżności dokładności dwóch serii pomiarów.

Przy założeniu, że populacje generalne mają rozkład odpowiednio: $\mathcal{N}(\mu_1, \sigma_1)$ i $\mathcal{N}(\mu_2, \sigma_2)$ i nie są znane parametry tych rozkładów, do weryfikacji hipotezy o równości wariancji wykorzystywana jest zmienna o rozkładzie ***F* Snedecora (Fishera)** o $v_1 = n_1 - 1$ i $v_2 = n_2 - 1$ (n_1, n_2 – liczebności prób wylosowanych z obydwu populacji) stopniach swobody:

$$F = \frac{s_1^2}{s_2^2}, \qquad gdzie: \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

Uwagi:

- Oznaczenia populacji przyjmowane są w taki sposób, że: $s_1^2 > s_2^2$.
- Hipoteza alternatywna H_1 (wobec hipotezy zerowej $H_0 : \sigma_1^2 = \sigma_2^2$) formułowana jest w postaci: $H_1 : \sigma_1^2 > \sigma_2^2$.

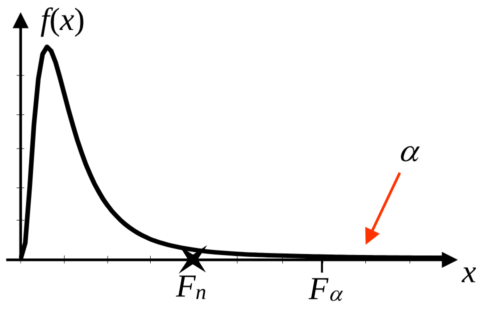
Przykład 6.

Wykonano dwie serie pomiarów długości detalu uzyskując wyniki takiej jak w przykładzie 3. Na poziomie istotności $\alpha = 0.01$ zweryfikować hipotezę o jednakowej wariancji obydwu serii pomiarów.

Wartość statystyki testowej wynosi ($s_1 \approx 2.00,\ s_2 \approx 1.00$):

$$F_n = \frac{s_1^2}{s_2^2} \approx \frac{2^2}{1^2} = 4.$$

Obszar krytyczny wyznacza się korzystając z odwrotności dystrybuanty rozkładu F o $v_1 = n_1 - 1$ i $v_2 = n_2 - 1$ stopniach swobody:



$$F_\alpha = F^{-1}_{F(9,7)}(1-\alpha) == F^{-1}_{F(9,7)}(0.99) \approx 6.72.$$

Wartość statystyki poza obszarem krytycznym, nie można odrzucić hipotezy H_0.

Graniczny poziom istotności p–value wyznacza się korzystając z dystrybuanty F o $v_1 = n_1 - 1$ i $v_2 = n_2 - 1$ stopniach swobody:

$$p\text{–}value = 1 - F_{F(9,7)}(F_n) \approx 1 - 0.96 = 0.04$$

Ponieważ założony *poziom istotności* α jest mniejszy od obliczonej wartości *p–value* więc nie ma podstaw do odrzucenia hipotezy H_0.

4.6. Weryfikacja hipotez dla frakcji (wskaźnika struktury, procentu)

Podstawowe badania cech niemierzalnych dotyczą badań frakcji elementów wyróżnionych. Frakcja (procent, wskaźnik struktury) jest estymowana zależnością:

$$\hat{p} = \frac{m}{n},$$

gdzie: m – liczba elementów wyróżnionych w próbie o liczebności n.

W przypadku dużej próby $(n \geq 100)$ rozkład frakcji z próby jest zbieżny do $\mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$.

Do weryfikacji hipotezy o frakcji wykorzystywana jest unormowana zmienna o rozkładzie $\mathcal{N}(0, 1)$:

$$U = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}.$$

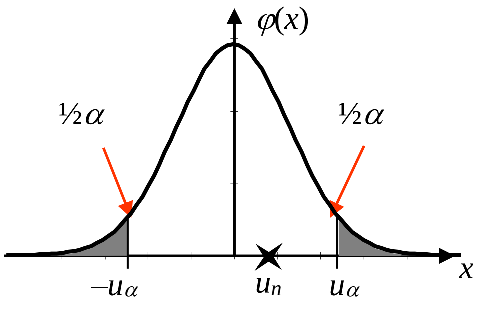
Przykład 7.

Wysunięto hipotezę, że wadliwość pewnego podzespołu wynosi 10%. W celu sprawdzenia tej hipotezy wylosowano próbkę 100 podzespołów i otrzymano w niej 15 podzespołów wadliwych. Zweryfikować hipotezę na poziomie istotności $\alpha = 0.05$.

1. Należy zweryfikować *hipotezę zerową* H_0: $p = 10\%$, wobec *hipotezy alternatywnej* H_1: $p \neq 10\%$.
2. Obliczona na podstawie wyników z próby wartość statystyki testowej (przy założeniu, że hipoteza H_0 jest prawdziwa) wynosi:

$$u_n = \frac{0.15 - 0.1}{\sqrt{\frac{0.1 \cdot 0.9}{100}}} \approx 1.67.$$

3. *Obszar krytyczny* wyznacza się korzystając z odwrotności dystrybuanty rozkładu $\mathcal{N}(0,1)$:

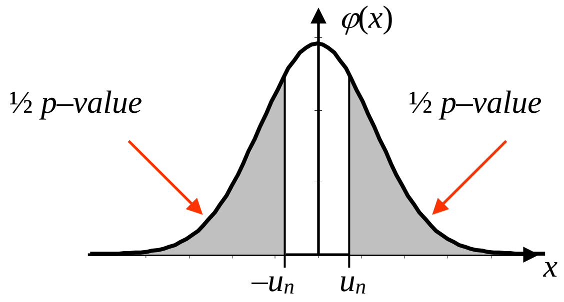


$$u_\alpha = -\Phi^{-1}\left(\frac{\alpha}{2}\right) = -\Phi^{-1}\left(\frac{0.05}{2}\right) \approx 1.96.$$

4. Ponieważ obliczona wartość statystyki testowej znajduje poza *obszarem krytycznym* ($|u_n| < |u_\alpha|$) więc nie ma podstaw do odrzucenia hipotezy H_0.

Identyczny rezultat otrzymuje się wyznaczając *graniczny poziom istotności p–value*.

3. *Graniczny poziom istotności p–value* wyznacza się korzystając z dystrybuanty rozkładu $\mathcal{N}(0,1)$:

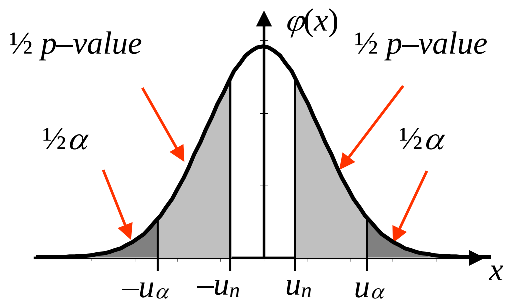


$$\tfrac{1}{2}\, p\text{–}value = \Phi(-u_n) = \Phi(-1.67) \approx 0.048$$

$$\Downarrow$$

$$p\text{–}value \approx 0.096$$

4. Ponieważ założony *poziom istotności* α jest mniejszy od obliczonej wartości *p–value* więc nie ma podstaw do odrzucenia hipotezy H_0.



4.7. Weryfikacja hipotez o równości frakcji dwóch populacji

Hipotezy dotyczące równości dwóch frakcji sprawdzane są w celu wyjaśnienia przypadkowej lub nieprzypadkowej rozbieżności frakcji elementów wyróżnionych w obu populacjach.
W przypadku gdy liczebności obydwu populacji są duże (≥100) do weryfikacji hipotezy o równości frakcji wykorzystywana jest zmienna o rozkładzie $\mathcal{N}(0, 1)$:

$$U = \frac{p_1 - p_2}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}},$$

gdzie: m_1 – liczba elementów wyróżnionych w próbie o liczebności n_1, m_2 – liczba elementów wyróżnionych w próbie o liczebności n_2, $p_1 = \frac{m_1}{n_1}$, $p_2 = \frac{m_2}{n_2}$, $\bar{p} = \frac{m_1 + m_2}{n_1 + n_2}$, $n = \frac{n_1 n_2}{n_1 + n_2}$.

Przykład 8.

Wysunięto hipotezę, że jakość produkcji pewnego wyrobu po wprowadzeniu nowej technologii nie uległa zmianie. Wylosowano 120 sztuk wyprodukowanych starą technologią i otrzymano 12 sztuk wadliwych, wśród wylosowanych 160 sztuk wyprodukowanych nową technologią i otrzymano 20 sztuk wadliwych.

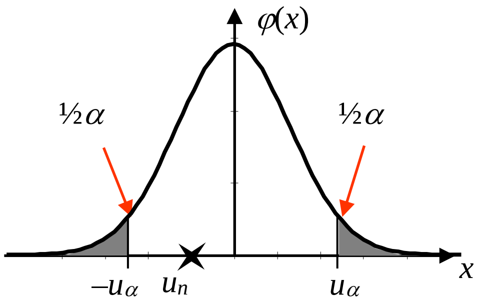
Zweryfikować na poziomie istotności $\alpha = 0.05$ hipotezę o jednakowym wskaźniku braków przy produkcji obydwoma metodami.

1. Należy zweryfikować *hipotezę zerową* H_0: $p_1 = p_2$, wobec *hipotezy alternatywnej* H_1: $p_1 \neq p_2$.
2. Obliczona na podstawie wyników z prób wartość statystyki testowej wynosi:

$p_1 = 0.1,\ p_2 = 0.125,\ \overline{p} \approx 0.11,\ n \approx 68.57$

$$u_n = \frac{0.1 - 0.125}{\sqrt{\frac{0.11 \cdot 0.89}{68.57}}} \approx -0.65.$$

3. *Obszar krytyczny* wyznacza się korzystając z odwrotności dystrybuanty rozkładu $\mathcal{N}(0,1)$:

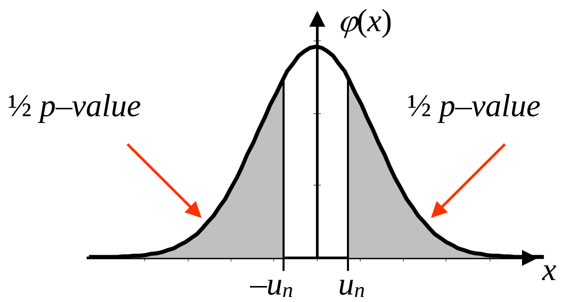


$$u_\alpha = \Phi^{-1}\left(\frac{\alpha}{2}\right) = \Phi^{-1}\left(\frac{0.05}{2}\right) \approx -1.96.$$

4. Ponieważ obliczona wartość statystyki testowej znajduje poza *obszarem krytycznym* ($|u_n| < |u_\alpha|$) więc nie ma podstaw do odrzucenia hipotezy H_0.

Identyczny rezultat otrzymuje się wyznaczając *graniczny poziom istotności p–value*.

3. *Graniczny poziom istotności p–value* wyznacza się korzystając z dystrybuanty rozkładu $\mathcal{N}(0,1)$:



$$\tfrac{1}{2}\ p\text{–}value = \Phi(-u_n) = \Phi(-0.65) \approx 0.26$$

$$\Downarrow$$

$$p\text{–}value \approx 0.52$$

4.8. Parametryczne testy istotności – podsumowanie

Lp.	Hipoteza H_0	Założenia	Statystyka testowa	Rozkład statystyki testowej
1.	$\mu = \mu_0$	znane σ lub nieznane σ i duża próba	$U = \frac{\bar{x} - \mu}{\sigma}\sqrt{n}$	$\mathcal{N}(0, 1)$ dla nieznanego σ: $\sigma \approx S$ lub $\sigma \approx s$
2.	$\mu = \mu_0$	nieznane σ i mała próba	$t = \frac{\bar{x} - \mu}{s}\sqrt{n}$	*t – Studenta* o (n – 1) stopniach swobody
3.	$\mu_1 = \mu_2$ *próby niezal.*	znane σ_1 i σ_2	$U = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$\mathcal{N}(0, 1)$
4.	$\mu_1 = \mu_2$ *próby niezal.*	nieznane ale równe σ_1 i σ_2	$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$, *gdzie:* $s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}}$	rozkład *t – Studenta* o ($n_1 - n_2$ – 2) stopniach swobody
5.	$\mu_1 = \mu_2$ *próby zal.*	znane σ_d	$U = \frac{\bar{d} - \mu_d}{\sigma_d}\sqrt{n}$	$\mathcal{N}(0, 1)$
6.	$\mu_1 = \mu_2$ *próby zal.*	nieznane σ_d	$t = \frac{\bar{d} - \mu_d}{s_d}\sqrt{n}$	*t – Studenta* o (n – 1) stopniach swobody
7.	$\sigma^2 = \sigma_0^2$		$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$	rozkład χ^2 o $v = n - 1$ stopniach swobody
8.	$\sigma_1^2 = \sigma_2^2$		$F = \frac{s_1^2}{s_2^2}, \quad (s_1^2 > s_2^2)$	rozkład F o $v_1 = n_1 - 1$ i $v_2 = n_2 - 1$ stopniach swobody
9.	$p = p_0$	duża próba $(n \geq 100)$	$U = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}}$.	$\mathcal{N}(0, 1)$
10.	$p_1 = p_2$	duża próba $(n_1, n_2 \geq 100)$	$U = \frac{p_1 - p_2}{\sqrt{\bar{p}(1-\bar{p})/n}}$, *gdzie:* $p_1 = m_1/n_1$, $p_2 = m_2/n_2$, $\bar{p} = (m_1 + m_2)/(n_1 + n_2)$, $n = n_1 n_2/(n_1 + n_2)$	$\mathcal{N}(0, 1)$