

Elementarna statystyka

Porównanie dwóch średnich

(Comparing two means)

Alexander Bendikov

Uniwersytet Wrocławski

21 maja 2016

W badaniu niedoborów żelaza u małych dzieci porównano wybrane losowo grupy dzieci, których matki wybrały różne sposoby karmienia. Jedna grupa dzieci była karmiona piersią. Druga grupa była karmiona standardowym mlekiem dla niemowląt, bez specjalnych suplementów żelazowych. W tabeli znajdują się poziomy hemoglobiny we krwi dzieci z obu grup w wieku 12 miesięcy.

Karmienie	n	$\bar{x}$	s
Naturalne	23	13,3	1,7
Odżywka	19	12,4	1,8

W badaniu niedoborów żelaza u małych dzieci porównano wybrane losowo grupy dzieci, których matki wybrały różne sposoby karmienia. Jedna grupa dzieci była karmiona piersią. Druga grupa była karmiona standardowym mlekiem dla niemowląt, bez specjalnych suplementów żelazowych. W tabeli znajdują się poziomy hemoglobiny we krwi dzieci z obu grup w wieku 12 miesięcy.

Karmienie	n	$\bar{x}$	s
Naturalne	23	13,3	1,7
Odżywka	19	12,4	1,8

- (a) Czy powyższe dane dostarczają istotnego dowodu na to, że średnie poziomy hemoglobiny (w całej populacji) są różne w tych dwóch grupach?

W badaniu niedoborów żelaza u małych dzieci porównano wybrane losowo grupy dzieci, których matki wybrały różne sposoby karmienia. Jedna grupa dzieci była karmiona piersią. Druga grupa była karmiona standardowym mlekiem dla niemowląt, bez specjalnych suplementów żelazowych. W tabeli znajdują się poziomy hemoglobiny we krwi dzieci z obu grup w wieku 12 miesięcy.

Karmienie	n	$\bar{x}$	s
Naturalne	23	13,3	1,7
Odżywka	19	12,4	1,8

- (a) Czy powyższe dane dostarczają istotnego dowodu na to, że średnie poziomy hemoglobiny (w całej populacji) są różne w tych dwóch grupach?
- (b) Jaki jest 95% przedział ufności dla średniej różnicy poziomu hemoglobiny w tych dwóch populacjach

Model matematyczny

1. Mamy dwie populacje, I oraz II, z każdej pobrano próbę losową SRS, odpowiednio X_1 i X_2

Model matematyczny

1. Mamy dwie populacje, I oraz II, z każdej pobrano próbę losową SRS, odpowiednio X_1 i X_2
2.
 - X_1 i X_2 są niezależne,
 - $\bar{x}_i \in N(\mu_i, \sigma_i/\sqrt{n_i})$, $i = 1, 2$.

Model matematyczny

1. Mamy dwie populacje, I oraz II, z każdej pobrano próbę losową SRS, odpowiednio X_1 i X_2
2.
 - X_1 i X_2 są niezależne,
 - $\bar{x}_i \in N(\mu_i, \sigma_i/\sqrt{n_i})$, $i = 1, 2$.
3. Dane: $X_1 : n_1, \bar{x}_1, s_1$, $X_2 : n_2, \bar{x}_2, s_2$.

Model matematyczny

1. Mamy dwie populacje, I oraz II, z każdej pobrano próbę losową SRS, odpowiednio X_1 i X_2
2.
 - X_1 i X_2 są niezależne,
 - $\bar{x}_i \in N(\mu_i, \sigma_i/\sqrt{n_i})$, $i = 1, 2$.
3. Dane: $X_1 : n_1, \bar{x}_1, s_1$, $X_2 : n_2, \bar{x}_2, s_2$.
4. Hipoteza: $H_0 : \mu_1 = \mu_2$
 $H_a : \mu_1 > \mu_2$,
 $H_a : \mu_1 < \mu_2$,
 $H_a : \mu_1 \neq \mu_2$.

Rozkład $\bar{x}_1 - \bar{x}_2$

- Średnia zmiennej $\bar{x}_1 - \bar{x}_2$ to $\mu_1 - \mu_2$,
- Wariancja zmiennej $\bar{x}_1 - \bar{x}_2$ to $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$,
- Rozkład zmiennej $\bar{x}_1 - \bar{x}_2$ jest normalny.

Rozkład $\bar{x}_1 - \bar{x}_2$

- Średnia zmiennej $\bar{x}_1 - \bar{x}_2$ to $\mu_1 - \mu_2$,
- Wariancja zmiennej $\bar{x}_1 - \bar{x}_2$ to $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$,
- Rozkład zmiennej $\bar{x}_1 - \bar{x}_2$ jest normalny.

z-statystyka dla 2 prób (σ_1, σ_2 znane):

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Rozkład $\bar{x}_1 - \bar{x}_2$

- Średnia zmiennej $\bar{x}_1 - \bar{x}_2$ to $\mu_1 - \mu_2$,
- Wariancja zmiennej $\bar{x}_1 - \bar{x}_2$ to $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$,
- Rozkład zmiennej $\bar{x}_1 - \bar{x}_2$ jest normalny.

z-statystyka dla 2 prób (σ_1, σ_2 znane):

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

t-statystyka dla 2 prób (σ_1, σ_2 nieznane):

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Uwaga: t -statystyka nie ma rozkładu t , ale można ją stosować, przyjmując liczbę stopni swobody

$$df = \min\{n_1 - 1, n_2 - 1\}.$$

Uwaga: t -statystyka nie ma rozkładu t , ale można ją stosować, przyjmując liczbę stopni swobody

$$df = \min\{n_1 - 1, n_2 - 1\}.$$

1. Przedziały ufności dla $\mu_1 - \mu_2$ (σ_1, σ_2 nieznane)

$$(\bar{x}_1 - \bar{x}_2) \pm t_* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

gdzie t^* jest górną wartością krytyczną dla $(1 - C)/2$ i liczbą stopni swobody $df = \min\{n_1 - 1, n_2 - 1\}$.

Uwaga: t -statystyka nie ma rozkładu t , ale można ją stosować, przyjmując liczbę stopni swobody

$$df = \min\{n_1 - 1, n_2 - 1\}.$$

1. Przedziały ufności dla $\mu_1 - \mu_2$ (σ_1, σ_2 nieznane)

$$(\bar{x}_1 - \bar{x}_2) \pm t_* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

gdzie t^* jest górną wartością krytyczną dla $(1 - C)/2$ i liczbą stopni swobody $df = \min\{n_1 - 1, n_2 - 1\}$.

2. Testowanie hipotez $H_0 : \mu_1 = \mu_2$, $H_a : \mu_1 \neq \mu_2$

Uwaga: t -statystyka nie ma rozkładu t , ale można ją stosować, przyjmując liczbę stopni swobody

$$df = \min\{n_1 - 1, n_2 - 1\}.$$

1. Przedziały ufności dla $\mu_1 - \mu_2$ (σ_1, σ_2 nieznane)

$$(\bar{x}_1 - \bar{x}_2) \pm t_* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

gdzie t^* jest górną wartością krytyczną dla $(1 - C)/2$ i liczbą stopni swobody $df = \min\{n_1 - 1, n_2 - 1\}$.

2. Testowanie hipotez $H_0 : \mu_1 = \mu_2$, $H_a : \mu_1 \neq \mu_2$

2.1 Obliczamy t -wartość przy $\mu_1 = \mu_2$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}},$$

Uwaga: t -statystyka nie ma rozkładu t , ale można ją stosować, przyjmując liczbę stopni swobody

$$df = \min\{n_1 - 1, n_2 - 1\}.$$

1. Przedziały ufności dla $\mu_1 - \mu_2$ (σ_1, σ_2 nieznane)

$$(\bar{x}_1 - \bar{x}_2) \pm t_* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

gdzie t^* jest górną wartością krytyczną dla $(1 - C)/2$ i liczbą stopni swobody $df = \min\{n_1 - 1, n_2 - 1\}$.

2. Testowanie hipotez $H_0 : \mu_1 = \mu_2$, $H_a : \mu_1 \neq \mu_2$

2.1 Obliczamy t -wartość przy $\mu_1 = \mu_2$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}},$$

2.2 Znajdujemy p -wartość, korzystając na przykład z tablic,

Uwaga: t -statystyka nie ma rozkładu t , ale można ją stosować, przyjmując liczbę stopni swobody

$$df = \min\{n_1 - 1, n_2 - 1\}.$$

1. Przedziały ufności dla $\mu_1 - \mu_2$ (σ_1, σ_2 nieznane)

$$(\bar{x}_1 - \bar{x}_2) \pm t_* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

gdzie t^* jest górną wartością krytyczną dla $(1 - C)/2$ i liczbą stopni swobody $df = \min\{n_1 - 1, n_2 - 1\}$.

2. Testowanie hipotez $H_0 : \mu_1 = \mu_2$, $H_a : \mu_1 \neq \mu_2$

2.1 Obliczamy t -wartość przy $\mu_1 = \mu_2$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}},$$

2.2 Znajdujemy p -wartość, korzystając na przykład z tablic,

2.3 Dochodzimy do konkluzji porównując p -wartość z założonym poziomem istotności testu α

Przykład: Poziomy hemoglobiny:

Karmienie	n	$\bar{x}$	s
Naturalne	23	13,3	1,7
Odżywka	19	12,4	1,8

1. Czy dane dostarczają istotnego statystycznie dowodu na to, że poziomy hemoglobiny w powyższych 2 grupach są różne?

Przykład: Poziomy hemoglobiny:

Karmienie	n	$\bar{x}$	s
Naturalne	23	13,3	1,7
Odżywka	19	12,4	1,8

1. Czy dane dostarczają istotnego statystycznie dowodu na to, że poziomy hemoglobiny w powyższych 2 grupach są różne?

1.1 $H_0 : \mu_1 = \mu_2, H_a : \mu_1 \neq \mu_2,$

1.2 Obliczamy t -wartość:

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = (13,3 - 12,4) / \sqrt{\frac{1,7^2}{23} + \frac{1,8^2}{19}} \simeq 1,654,$$

1.3 p -wartość: $df = \min\{22, 18\} = 18, p = 0,10, p = 0,05$ skąd

$$t_* = 1,734, t_* = 1,734$$

1.4 Konkluzja: Dane nie dostarczają statystycznego dowodu na to, że $\mu_1 \neq \mu_2$ (czyli nie odrzucamy $H_0 : \mu_1 = \mu_2$) na poziomie istotności $\alpha = 2 \times 0,05 = 0,1$.

2. Obliczmy 95% przedział ufności

$$(\bar{x}_1 - \bar{x}_2) \pm t_* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (\bar{x}_1 - \bar{x}_2) \pm m.$$

Korzystając, na przykład z tablic, dla $df = 18$, $p = (1 - C)/2 = 0,025$, $t_* = 2,101$. Otrzymujemy przedział ufności

$$\begin{aligned} (13,3 - 12,4) \pm 2,101 \left(\frac{1,7^2}{23} + \frac{1,8^2}{19} \right)^{1/2} &= 0,9 \pm 2,101 \cdot 0,544 = \\ &= 0,9 \pm 1,143; \\ -0,243 &\leq \mu_1 - \mu_2 \leq 2,043. \end{aligned}$$

Uwagi: 1. 95% przedział ufności zawiera wartość $\mu_1 - \mu_2 = 0$. W takim razie nie odrzucamy hipotezy $H_0 : \mu_1 = \mu_2$, na poziomie istotności 5%.

Uwagi: 1. 95% przedział ufności zawiera wartość $\mu_1 - \mu_2 = 0$. W takim razie nie odrzucamy hipotezy $H_0 : \mu_1 = \mu_2$, na poziomie istotności 5%.

2. Powyższy wynik jest dokładniejszy niż poprzedni (testowanie hipotezy). Przynajmniej dla hipotezy H_a dwustronnej przedział ufności niesie więcej informacji.

Wnioskowanie o proporcjach w populacji

Przykłady: 1. C. Buffon (1707-1788) rzucił monetą 4040 razy, i otrzymał 2048 orłów, $\hat{p} = 0,5069$.

Wnioskowanie o proporcjach w populacji

Przykłady: 1. C. Buffon (1707-1788) rzucił monetą 4040 razy, i otrzymał 2048 orłów, $\hat{p} = 0,5069$.

2. K. Pearson (1857-1936) rzucił monetą 24000 razy i otrzymał 12012 orłów, $\hat{p} = 0,5005$.

Wnioskowanie o proporcjach w populacji

Przykłady: 1. C. Buffon (1707-1788) rzucił monetą 4040 razy, i otrzymał 2048 orłów, $\hat{p} = 0,5069$.

2. K. Pearson (1857-1936) rzucił monetą 24000 razy i otrzymał 12012 orłów, $\hat{p} = 0,5005$.

- Czy te wyniki są wystarczającym dowodem na to, że używane monety nie były symetryczne?
- Jaki jest 95% przedział ufności dla średniej p w obu przypadkach.

Wnioskowanie o proporcjach w populacji

Przykłady: 1. C. Buffon (1707-1788) rzucił monetą 4040 razy, i otrzymał 2048 orłów, $\hat{p} = 0,5069$.

2. K. Pearson (1857-1936) rzucił monetą 24000 razy i otrzymał 12012 orłów, $\hat{p} = 0,5005$.

- Czy te wyniki są wystarczającym dowodem na to, że używane monety nie były symetryczne?
- Jaki jest 95% przedział ufności dla średniej p w obu przypadkach.

Model: Interesuje nas pewna własność, którą osobniki w dużej populacji mogą mieć lub nie mieć. Proporcja w populacji to

$$p = \frac{\# \text{ osobników mających własność}}{\# \text{ wszystkich osobników w populacji}}.$$

Wnioskowanie o proporcjach w populacji

Przykłady: 1. C. Buffon (1707-1788) rzucił monetą 4040 razy, i otrzymał 2048 orłów, $\hat{p} = 0,5069$.

2. K. Pearson (1857-1936) rzucił monetą 24000 razy i otrzymał 12012 orłów, $\hat{p} = 0,5005$.

- Czy te wyniki są wystarczającym dowodem na to, że używane monety nie były symetryczne?
- Jaki jest 95% przedział ufności dla średniej p w obu przypadkach.

Model: Interesuje nas pewna własność, którą osobniki w dużej populacji mogą mieć lub nie mieć. Proporcja w populacji to

$$p = \frac{\# \text{ osobników mających własność}}{\# \text{ wszystkich osobników w populacji}}.$$

Ogólnie proporcja p nie jest znana. Wybieramy próbę i obliczamy proporcję w próbie:

$$\hat{p} = \frac{\# \text{ osobników w próbie mających własność}}{\# \text{ wszystkich osobników w próbie}}.$$

Główne założenie: Cała populacja jest co najmniej 10 razy większa niż próba.

Główne założenie: Cała populacja jest co najmniej 10 razy większa niż próba.

Konkluzja: # osobników w próbie mających daną własność ma rozkład który jest bliski rozkładowi dwumianowemu.

Dlatego, jeżeli wielkość próby to n , to mamy

$$E(\hat{p}) = p, \quad SD(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}.$$

Główne założenie: Cała populacja jest co najmniej 10 razy większa niż próba.

Konkluzja: # osobników w próbie mających daną własność ma rozkład który jest bliski rozkładowi dwumianowemu.

Dlatego, jeżeli wielkość próby to n , to mamy

$$E(\hat{p}) = p, \quad SD(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}.$$

Testowanie hipotezy $H_0 : p = p_0$, $H_a : p \neq p_0$ (lub $p \leq p_0$). Jeżeli n , np_0 , $n(1-p_0)$ są wszystkie ≥ 10 , możemy zastosować Centralne Twierdzenie Graniczne: rozkład zmiennej

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

jest zbliżony do normalnego $N(0, 1)$.

Przykłady:

1. (Moneta Buffona): $H_0 : p = 0,5$, $H_a : p \neq 0,5$

1.1 $z\text{-wartość} = \frac{0,5069 - 0,5}{\sqrt{\frac{0,5 \cdot 0,5}{4040}}} = 0,88$

1.2 $p\text{-wartość} = 2 \times 0,1894 = 0,3788$

Przykłady:

1. (Moneta Buffona): $H_0 : p = 0,5$, $H_a : p \neq 0,5$

1.1 $z\text{-wartość} = \frac{0,5069 - 0,5}{\sqrt{\frac{0,5 \cdot 0,5}{4040}}} = 0,88$

1.2 $p\text{-wartość} = 2 \times 0,1894 = 0,3788$

2. (Moneta Pearsona): $H_0 : p = 0,5$, $H_a : p \neq 0,5$

2.1 $z\text{-wartość} = \frac{0,50005 - 0,5}{\sqrt{\frac{0,5 \cdot 0,5}{24000}}} = 0,155$

2.2 $p\text{-wartość} = 2 \times 0,4364 = 0,8728$

Przykłady:

1. (Moneta Buffona): $H_0 : p = 0,5$, $H_a : p \neq 0,5$

1.1 $z\text{-wartość} = \frac{0,5069 - 0,5}{\sqrt{\frac{0,5 \cdot 0,5}{4040}}} = 0,88$

1.2 $p\text{-wartość} = 2 \times 0,1894 = 0,3788$

2. (Moneta Pearsona): $H_0 : p = 0,5$, $H_a : p \neq 0,5$

2.1 $z\text{-wartość} = \frac{0,50005 - 0,5}{\sqrt{\frac{0,5 \cdot 0,5}{24000}}} = 0,155$

2.2 $p\text{-wartość} = 2 \times 0,4364 = 0,8728$

Konkluzja: Ani wyniki Buffona ani Pearsona nie sugerują, że ich monety były niesymetryczne.

Przykłady:

1. (Moneta Buffona): $H_0 : p = 0,5$, $H_a : p \neq 0,5$

1.1 $z\text{-wartość} = \frac{0,5069 - 0,5}{\sqrt{\frac{0,5 \cdot 0,5}{4040}}} = 0,88$

1.2 $p\text{-wartość} = 2 \times 0,1894 = 0,3788$

2. (Moneta Pearsona): $H_0 : p = 0,5$, $H_a : p \neq 0,5$

2.1 $z\text{-wartość} = \frac{0,50005 - 0,5}{\sqrt{\frac{0,5 \cdot 0,5}{24000}}} = 0,155$

2.2 $p\text{-wartość} = 2 \times 0,4364 = 0,8728$

Konkluzja: Ani wyniki Buffona ani Pearsona nie sugerują, że ich monety były niesymetryczne.

Przedziały ufności: Przedział ufności dla proporcji p to $\hat{p} \pm z_* SE$,

$$SE \approx \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

Przykłady:

1. (Moneta Buffona): $H_0 : p = 0,5$, $H_a : p \neq 0,5$

1.1 $z\text{-wartość} = \frac{0,5069 - 0,5}{\sqrt{\frac{0,5 \cdot 0,5}{4040}}} = 0,88$

1.2 $p\text{-wartość} = 2 \times 0,1894 = 0,3788$

2. (Moneta Pearsona): $H_0 : p = 0,5$, $H_a : p \neq 0,5$

2.1 $z\text{-wartość} = \frac{0,50005 - 0,5}{\sqrt{\frac{0,5 \cdot 0,5}{24000}}} = 0,155$

2.2 $p\text{-wartość} = 2 \times 0,4364 = 0,8728$

Konkluzja: Ani wyniki Buffona ani Pearsona nie sugerują, że ich monety były niesymetryczne.

Przedziały ufności: Przedział ufności dla proporcji p to $\hat{p} \pm z_* SE$,

$$SE \approx \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

Przykłady: 1. 95% przedział ufności dla monety Buffona:

$$\begin{aligned}\hat{p} \pm z_* SE &\approx 0,5069 \pm 1,960 \sqrt{\frac{0,5069 \cdot 0,4931}{4040}} = \\ &= 0,5069 \pm 0,0154, \\ 0,4915 &< p < 0,5223.\end{aligned}$$

2. 99% przedział ufności dla monety Pearsona:

$$\begin{aligned}\hat{p} \pm z_* SE &\approx 0,5005 \pm 2,576 \sqrt{\frac{0,5005 \cdot 0,4995}{24000}} = \\ &= 0,5005 \pm 0,0083, \\ 0,4922 &< p < 0,5088.\end{aligned}$$

2. 99% przedział ufności dla monety Pearsona:

$$\begin{aligned}\hat{p} \pm z_* SE &\approx 0,5005 \pm 2,576 \sqrt{\frac{0,5005 \cdot 0,4995}{24000}} = \\ &= 0,5005 \pm 0,0083, \\ 0,4922 &< p < 0,5088.\end{aligned}$$

Rozmiar próby dla założonego marginesu błędu

- C – poziom ufności,
- m – założony margines błędu
- p^* – przypuszczalna wartość proporcji

$$n = \left(\frac{z_*}{m} \right)^2 p^*(1 - p^*).$$

(z^* jest uzyskane z tablic dla założonego C .)

Przykład: PTC jest substancją chemiczną, którą niektórzy ludzie odbierają jako bardzo gorzką, a inni ludzie odbierają jako pozbawioną smaku.

Zdolność do odczuwania goryczy PTC jest dziedziczna. Około 75% Włochów posiada tę zdolność. Chcemy oszacować tę proporcję wśród Amerykanów, którzy wśród swoich dziadków (rodziców rodziców) mają co najmniej jednego Włocha/jedną Włoszkę. Jak duża powinna być próba, aby oszacować z marginesem błędu $m = \pm 0,04$ na 95% poziomie ufności? Rozwiązanie: Nasze przypuszczenie odnośnie proporcji to $p^* = 0,75$. Dla 95% poziomu ufności mamy $z_* = 1,96$. Dla osiągnięcia $m = 0,04$ wielkość próby musi wynieść co najmniej

$$n = \left(\frac{z_*}{m} \right)^2 p^*(1 - p^*) = \left(\frac{1,96}{0,04} \right)^2 0,75 \cdot 0,25 = 450,1875.$$

Przyjmujemy $n = 451$.