Weryfikacja hipotez statystycznych

Przykład. Producent pewnych detali twierdzi, że wadliwość jego produkcji nie przekracza 2%. Odbiorca pewnej partii tego produktu chce sprawdzić, czy może wierzyć producentowi. W jaki sposób ma to zrobić?

Krok 1. Zakładamy, że partia ma wadliwość 2%.

Krok 2. Pobierana jest próba elementów z partii towaru (np. 100 elementów).

k	$P\{X = k\}$	$P\{X \geq k\}$
0	0.135335	1.000000
1	0.270671	0.864665
2	0.270670	0.593994
3	0.180447	0.323324
4	0.090224	0.142877
5	0.036089	0.052653
6	0.012030	0.016564
7	0.004297	0.004534
8	0.000191	0.000237

Krok 3 (wnioskowanie).

Zaobserwowano $k = 7$ wadliwych:

1. Przypuszczenie jest słuszne i próba „pechowa”
lub
2. Próba jest „dobra”, a przypuszczenie złe.

Uznać twierdzenie producenta za nieprawdziwe!

Zaobserwowano co najmniej siedem wadliwych
Wnioski jak wyżej

Ostatecznie:

Po zaobserwowaniu więcej niż sześciu wadliwych elementów raczej uznać twierdzenie producenta za nieprawdziwe.
W przeciwnym przypadku można uznać twierdzenie producenta za uzasadnione.

Hipotezą statystyczną nazywamy dowolne przypuszczenie dotyczące rozkładu prawdopodobieństwa cechy w populacji.
Oznaczenie H_0

Testem hipotezy statystycznej nazywamy postępowanie mające na celu odrzucenie lub nie odrzucenie hipotezy statystycznej.

Statystyką testową nazywamy funkcję próby na podstawie której wnioskuje się o odrzuceniu lub nie hipotezy statystycznej.

Rzeczywistość: hipoteza H_0	Wniosek o hipotezie H_0 nie odrzucać	odrzucić
prawdziwa	prawidłowy	nieprawidłowy
nieprawdziwa	nieprawidłowy	prawidłowy

Błędem I rodzaju nazywamy błąd wnioskowania polegający na odrzuceniu hipotezy, gdy w rzeczywistości jest ona prawdziwa.

Błędem II rodzaju nazywamy błąd wnioskowania polegający na nieodrzuceniu hipotezy, gdy w rzeczywistości jest ona fałszywa.

Poziomem istotności nazywamy dowolną liczbę z przedziału $(0, 1)$ określającą prawdopodobieństwo popełnienia błędu I rodzaju.
Oznaczenie: α

Mocą testu nazywamy prawdopodobieństwo odrzucenia testowanej hipotezy, gdy jest ona nieprawdziwa, czyli prawdopodobieństwo nie popełnienia błędu II rodzaju.
Oznaczenie: $1 - \beta$

Rozkład normalny
Porównanie z normą

$$H_0 : \mu = \mu_0$$

Cecha X ma rozkład normalny $N(\mu, \sigma^2)$
Średnia μ oraz wariancja σ^2 są nieznane

Test Studenta (poziom istotności α)

Próba: $X_1, \ldots, X_n$
Statystyka testowa

$$t_{\text{emp}} = \frac{\bar{X} - \mu_0}{S}\sqrt{n} \ .$$

Wartość krytyczna $t(\alpha; n-1)$

Jeżeli $|t_{\text{emp}}| > t(\alpha; n-1)$, to hipotezę $H_0 : \mu = \mu_0$ odrzucamy.

Przykład. Przypuszczenie: maszyna pakująca kostki masła nastawiona na jednostkową masę 250 g uległa po pewnym czasie rozregulowaniu. W celu weryfikacji tego przypuszczenia z bieżącej produkcji pobrano próbę otrzymując wyniki 254, 269, 254, 248, 263, 256, 258, 261, 264, 258. Czy można na tej podstawie sądzić, że maszyna uległa rozregulowaniu?

Populacja:
paczkowane kostki masła

Cecha X:
masa kostki masła

Założenie:
cecha X ma rozkład normalny $N(\mu, \sigma^2)$

Formalizacja:
Rozregulowanie maszyny może być interpretowane jako odejście od nominalnej wagi. Zatem należy zbadać, czy średnia μ wynosi 250, czyli weryfikujemy hipotezę $H_0 : \mu = 250$

Technika statystyczna:
test Studenta (test t)
poziom istotności $\alpha = 0.05$

Obliczenia

$$\bar{x} = 258.5,\ s^2 = 36.05,\ t_{\mathrm{emp}} = 4.47$$

Wartość krytyczna: $t(0.05; 9) = 2.2622$

Odpowiedź: hipotezę odrzucamy

Wniosek: maszyna uległa rozregulowaniu

Moc testu

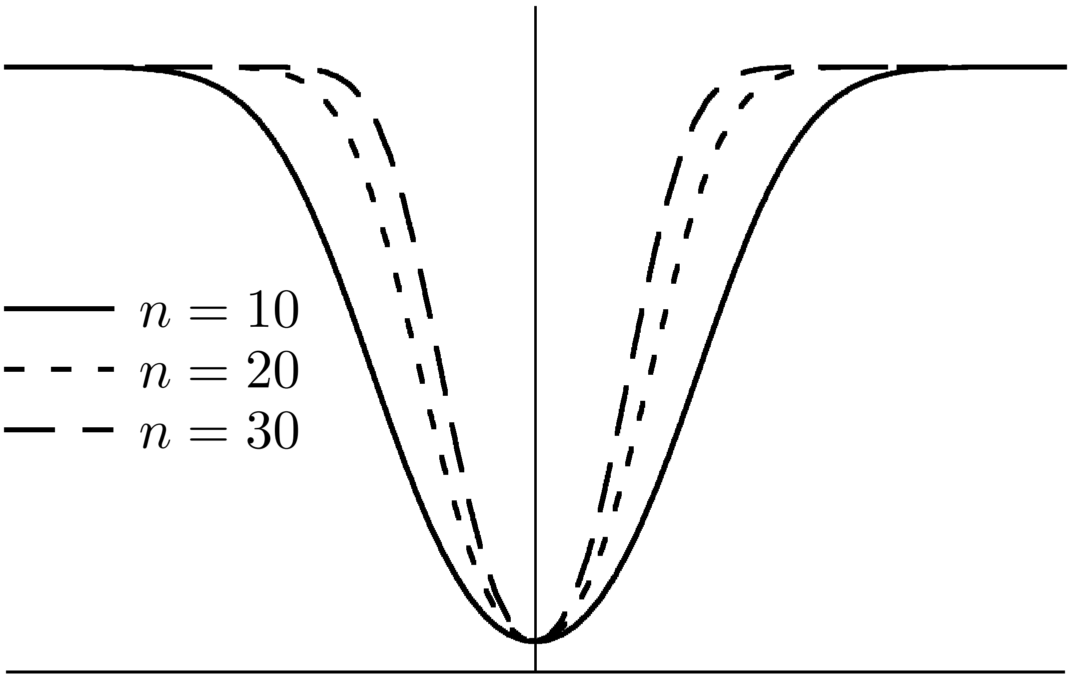
$$\text{Moc testu} = 1 - P\{\text{błąd II rodzaju}\}$$

$$\text{Moc testu} = P\{\text{odrzucenie nieprawdziwej } H_0\}$$

Moc testu Studenta hipotezy $H_0 : \mu = \mu_0$

$$\mathcal{M}(\mu) = P\{|t_{\text{emp}}| > t(\alpha; n-1) | X \sim N(\mu, \sigma^2)\}$$

$$\mathcal{M}(\mu_0) = \alpha$$



Przedział ufności a test hipotezy $H_0 : \mu = \mu_0$

Cecha $X \sim N(\mu, \sigma^2)$

$$H_0 : \mu = \mu_0$$

H_0 nie odrzucamy na poziomie istotności α

$$\Updownarrow$$

$$|t_{\text{emp}}| < t(\alpha; n-1)$$

$$\Updownarrow$$

$$-t(\alpha; n-1) < \frac{\bar{X} - \mu_0}{S}\sqrt{n} < t(\alpha; n-1)$$

$$\Updownarrow$$

$$\mu_0 \in \left(\bar{X} - t(\alpha; n-1)\frac{S}{\sqrt{n}}, \bar{X} + t(\alpha; n-1)\frac{S}{\sqrt{n}}\right)$$

$$\Updownarrow$$

μ_0 należy do przedziału ufności

na poziomie ufności $1 - \alpha$

$$H_0 : \mu \leq \mu_0$$

Cecha X ma rozkład normalny $N(\mu, \sigma^2)$
Średnia μ oraz wariancja σ^2 są nieznane

Test Studenta (poziom istotności α)

Próba: $X_1, \ldots, X_n$
Statystyka testowa

$$t_{\mathrm{emp}} = \frac{\bar{X} - \mu_0}{S}\sqrt{n}\ .$$

Wartość krytyczna $t(2\alpha; n-1)$

Jeżeli $t_{\mathrm{emp}} > t(2\alpha; n-1)$, to hipotezę $H_0 : \mu \leq \mu_0$ odrzucamy.

$$H_0 : \sigma^2 = \sigma_0^2$$

Cecha X ma rozkład normalny $N(\mu, \sigma^2)$
Średnia μ oraz wariancja σ^2 są nieznane

Test chi–kwadrat (poziom istotności α)

Próba: $X_1, \ldots, X_n$
Statystyka testowa

$$\chi^2_{\text{emp}} = \frac{\text{var}X}{\sigma_0^2}$$

Wartości krytyczne
$\chi^2\left(1 - \frac{\alpha}{2}; n-1\right)$ oraz $\chi^2\left(\frac{\alpha}{2}; n-1\right)$

Jeżeli

$\chi^2_{\text{emp}} < \chi^2\left(1 - \frac{\alpha}{2}; n-1\right)$
lub
$\chi^2_{\text{emp}} > \chi^2\left(\frac{\alpha}{2}; n-1\right)$,

to hipotezę $H_0 : \sigma^2 = \sigma_0^2$ odrzucamy.

$$H_0 : \sigma^2 \leq \sigma_0^2$$

Cecha X ma rozkład normalny $N(\mu, \sigma^2)$
Średnia μ oraz wariancja σ^2 są nieznane

Test chi–kwadrat (poziom istotności α)

Próba: $X_1, \ldots, X_n$
Statystyka testowa

$$\chi^2_{\text{emp}} = \frac{\text{var}X}{\sigma_0^2}$$

Wartość krytyczna $\chi^2\,(\alpha; n-1)$

Jeżeli $\chi^2_{\text{emp}} > \chi^2\,(\alpha; n-1)$, to hipotezę $H_0 : \sigma^2 \leq \sigma_0^2$ odrzucamy.

Przykład. Na podstawie obserwacji prowadzonych przez długi okres czasu stwierdzono, że dzienny udój uzyskiwany w pewnym stadzie krów jest wielkością losową, zaś przeciętny dzienny udój mleka wyraża sie liczbą z przedziału $(900, 1200)$. Rachunek finansowy pokazał, że produkcja mleka jest opłacalna, jeżeli całkowity dzienny udój będzie wynosił nie mniej niż $d = 700\, l$ mleka przez co najmniej 280 dni w roku. W jaki sposób można zbadać, czy produkcja mleka jest opłacalna?

Populacja:

Cecha:
całkowity dzienny udój

Założenia:
Cecha X ma rozkład $N(\mu, \sigma^2)$
$\mu_d = 900 \leq \mu \leq \mu_g = 1200$

Formalizacja problemu

$$P\{X \geq d\} \geq p = \frac{280}{350}$$

$$P\{X \geq d\} = 1 - F\left(\frac{d-\mu}{\sigma}\right) \geq 1 - F\left(\frac{d-\mu_d}{\sigma}\right)$$

$$1 - F\left(\frac{d-\mu_d}{\sigma}\right) \geq 1 - p \Rightarrow F\left(\frac{d-\mu_d}{\sigma}\right) \leq 1 - p$$

$$\frac{d-\mu_d}{\sigma} \leq F^{-1}(1-p) = u_{1-p}$$

d, μ_d oraz p są ustalone, więc

$$\sigma^2 \geq \sigma_0^2 = \left(\frac{d-\mu_d}{u_{1-p}}\right)^2 = 56472$$

Produkcja mleka jest opłacalna, jeżeli wariancja σ^2 dziennych udojów jest większa niż $\sigma_0^2 = 56472$.

$$H_0 : \sigma^2 \leq 56472$$

Rozkład dwupunktowy
Porównanie z normą

$$H_0 : p = p_0$$

Cecha X ma rozkład $D(p)$
Próba: $X_1, \ldots, X_n$ ($X_i = 0$ lub $= 1$)

Statystyka testowa

$$Y = \sum_{i=1}^{n} X_i$$

Jeżeli $Y \leq k_1$ lub $Y \geq k_2$, to hipotezę $H_0 : p = p_0$ należy odrzucić.

Liczby k_1 oraz k_2 dobrane są tak, że jeżeli Y jest zmienną losową o rozkładzie $B(n, p_0)$, to

$$P\{Y \leq k_1 \text{ lub } Y \geq k_2\} \leq \alpha$$

$$H_0 : p = p_0$$

Test przybliżony (poziom istotności α)
Przypadek: n „duże”

Statystyka testowa

$$u_{\mathrm{emp}} = \frac{Y - np_0}{\sqrt{np_0(1-p_0)}}$$

Wartość krytyczna $u_{1-\alpha/2}$

Jeżeli $|u_{\mathrm{emp}}| > u_{1-\alpha/2}$, to $H_0 : p = p_0$ odrzucamy

$$H_0 : p \leq p_0$$

Test przybliżony (poziom istotności α)
Przypadek: n „duże”

Statystyka testowa

$$u_{\mathrm{emp}} = \frac{Y - np_0}{\sqrt{np_0(1 - p_0)}}$$

Wartość krytyczna $u_{1-\alpha}$

Jeżeli $u_{\mathrm{emp}} > u_{1-\alpha}$, to $H_0 : p \leq p_0$ odrzucamy

Przykład. W swojej ofercie sprzedaży stawu rybnego jego właściciel podaje, iż w stawie żyje co najmniej tysiąc karpi. Potencjalny nabywca zainteresowany jest sprawdzeniem prawdziwości tego twierdzenia. W tym celu wyłowiono sto karpi i po zaobrączkowaniu ich wpuszczono je z powrotem do stawu. Po jakimś czasie ponownie odłowiono sto ryb i stwierdzono, że wśród nich jest piętnaście zaobrączkowanych. Czy w świetle uzyskanych wyników można reklamę uznać za prawdziwą?

Populacja:
ryby w stawie

Cecha:
zaobrączkowanie ryby

Założenia:
Cecha X ma rozkład $D(p)$

Formalizacja problemu
Jeżeli w stawie żyje co najmniej N ryb, to odsetek zaobrączkowanych jest co najwyżej $100/N$. Zgodnie z twierdzeniem właściciela, $N \geq 1000$, czyli odsetek ryb zaobrączkowanych nie przekracza 0.1.

Technika statystyczna
Przybliżony test hipotezy $H_0 : p \leq 0.1$
Poziom istotności: $\alpha = 0.05$

Obliczenia

$$Y = 15 \quad n = 100$$

$$u_{\text{emp}} = \frac{Y - np_0}{\sqrt{np_0(1-p_0)}} = \frac{15-10}{\sqrt{100 \cdot 0.1 \cdot 0.9}} = 1.6667$$

Wartość krytyczna: $u_{1-0.05} = 1.6449$

Odpowiedź: hipotezę odrzucamy

Wniosek: należy uznać, że ogólna liczb ryb w stawie jest mniejsza niż podana w ofercie

Porównanie dwóch rozkładów normalnych

Założenia:
1. $X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2)$
2. X_1, X_2 są niezależne

Czy $\mu_1 = \mu_2$?

Czy $\sigma_1^2 = \sigma_2^2$?

Próby: $X_{11}, \ldots, X_{1n_1}$; $X_{21}, \ldots, X_{2n_2}$

$$\bar{X}_1, \quad \mathrm{var}X_1, \quad s_1^2 = \frac{\mathrm{var}X_1}{n_1 - 1}$$

$$\bar{X}_2, \quad \mathrm{var}X_2, \quad s_2^2 = \frac{\mathrm{var}X_2}{n_2 - 1}$$

$$H_0 : \mu_1 = \mu_2$$

Założenie $\sigma_1^2 = \sigma_2^2$

Test Studenta (poziom istotności α)

Statystyka testowa

$$t_{\mathrm{emp}} = \frac{\bar{X}_1 - \bar{X}_2}{S_r}$$

$$S_r = \sqrt{S_e^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}, \quad S_e^2 = \frac{\mathrm{var} X_1 + \mathrm{var} X_2}{n_1 + n_2 - 2}$$

Wartość krytyczna $t(\alpha; n_1 + n_2 - 2)$

Jeżeli $|t_{\mathrm{emp}}| > t(\alpha; n_1 + n_2 - 2)$,
to hipotezę $H_0 : \mu_1 = \mu_2$ odrzucamy

$$H_0 : \mu_1 = \mu_2$$

Bez założenia $\sigma_1^2 = \sigma_2^2$

Test V Behrensa–Fishera (poziom istotności α)

Statystyka testowa

$$V = \frac{\bar{X}_1 - \bar{X}_2}{S_r}$$

$$S_r = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

Wartość krytyczna $V(\alpha; n_1 - 1, n_2 - 1, c)$ $(n_1 \leq n_2)$

$$c = \frac{S_1^2/n_1}{S_1^2/n_1 + S_2^2/n_2}$$

Jeżeli $|V| > V(\alpha; n_1 - 1, n_2 - 1, c)$,
to hipotezę $H_0 : \mu_1 = \mu_2$ odrzucamy

Przykład. Porównać przeciętne osiągnięcia punktowe pań i panów na klasówce ze statystyki
Panowie:

$$n_1 = 138, \quad \sum x_{1i} = 82.833, \quad \mathrm{var} x_1 = 1.65841$$

Panie:

$$n_2 = 162, \quad \sum x_{2i} = 93.733, \quad \mathrm{var} x_2 = 2.23348$$

Populacja 1:
Słuchacze podstawowego kursu statystyki
Populacja 2:
Słuchaczki podstawowego kursu statystyki

Cecha X: ilość punktów zdobytych na klasówce

Założenia:
cecha X ma w populacji 1 rozkład $N(\mu_1, \sigma_1^2)$
cecha X ma w populacji 2 rozkład $N(\mu_2, \sigma_2^2)$
$\sigma_1^2 = \sigma_2^2$

Zadanie: zweryfikować hipotezę $H_0 : \mu_1 = \mu_2$

Technika statystyczna:
test t
poziom istotności 0.05

Obliczenia

$$\bar{x}_1 = 0.60024 \quad \bar{x}_2 = 0.57860$$

$$s_r^2 = \frac{1.65841 + 2.23348}{138 + 162 - 2}\left(\frac{1}{138} + \frac{1}{162}\right)$$
$$= 0.000175255$$

$$t_{\text{emp}} = \frac{0.60024 - 0.57860}{\sqrt{0.000175255}} = 1.634$$

Wartość krytyczna $t(0.05; 298) \approx 1.96$

Odpowiedź: hipotezy nie odrzucamy

Wniosek.
Średnie ilości punktów uzyskiwane przez panie i panów można traktować jako porównywalne.

Przedział ufności a test hipotezy $H_0 : \mu_1 = \mu_2$

Cecha $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, $\sigma_1^2 = \sigma_2^2$

$$H_0 : \mu_1 = \mu_2$$

H_0 nie odrzucamy na poziomie istotności α

$$\Updownarrow$$

$$|t_{\mathrm{emp}}| < t(\alpha; n_1 + n_2 - 2)$$

$$\Updownarrow$$

$$-t(\alpha; n_1 + n_2 - 2) < \frac{\bar{X}_1 - \bar{X}_2}{S_r} < t(\alpha; n_1 + n_2 - 2)$$

$$\Updownarrow$$

$$0 \in \left(\bar{X}_1 - \bar{X}_2 \pm t(\alpha; n_1 + n_2 - 2) S_r\right)$$

$$\Updownarrow$$

0 należy do przedziału ufności

na poziomie ufności $1 - \alpha$

Przykład. Porównać wartości średnie dwóch cech X_1 oraz X_2 o rozkładach normalnych.

$$H_0 : \mu_1 = \mu_2$$

Test V Behrensa–Fishera ($\alpha = 0.05$)

Próby:

$$\begin{array}{lll} n_1 = 20 & \bar{x}_1 = 74.40 & s_1^2 = 15.41 \\ n_2 = 20 & \bar{x}_2 = 65.05 & s_2^2 = 83.73 \end{array}$$

$$V = \frac{74.40 - 65.05}{\sqrt{\frac{15.41}{20} + \frac{83.73}{20}}} = \frac{9.35}{\sqrt{4.96}} = 4.19$$

$$c = \frac{15.41/20}{15.41/20 + 83.73/20} = \frac{0.77}{4.96} = 0.15$$

Wartość krytyczna $V(0.05; 19, 19, 0.15) = 2.06$

Ponieważ $|V| > V(0.05; 19, 19, 0.15)$, więc hipotezę $H_0 : \mu_1 = \mu_2$ odrzucamy.

$$H_0 : \mu_1 \leq \mu_2$$

Założenie $\sigma_1^2 = \sigma_2^2$

Test Studenta (poziom istotności α)

Statystyka testowa

$$t_{\text{emp}} = \frac{\bar{X}_1 - \bar{X}_2}{S_r}$$

Wartość krytyczna $t(2\alpha; n_1 + n_2 - 2)$

Jeżeli $t_{\text{emp}} > t(2\alpha; n_1 + n_2 - 2)$, to hipotezę $H_0 : \mu_1 \leq \mu_2$ odrzucamy

....

Bez założenia $\sigma_1^2 = \sigma_2^2$

Test V Behrensa–Fishera (poziom istotności α)

Statystyka testowa

$$V = \frac{\bar{X}_1 - \bar{X}_2}{S_r}$$

Wartość krytyczna $V(2\alpha; n_1 - 1, n_2 - 1, c)$ $(n_1 \leq n_2)$

Jeżeli $V > V(2\alpha; n_1 - 1, n_2 - 1, c)$, to hipotezę $H_0 : \mu_1 \leq \mu_2$ odrzucamy.

$$H_0 : \sigma_1^2 = \sigma_2^2$$

Test F (poziom istotności α)

Statystyka testowa

$$F_{\text{emp}} = \frac{S_1^2}{S_2^2}$$

Wartości krytyczne

$F\left(1 - \dfrac{\alpha}{2}; n_1 - 1, n_2 - 1\right)$

$F\left(\dfrac{\alpha}{2}; n_1 - 1, n_2 - 1\right)$

Jeżeli

$F_{\text{emp}} < F\left(1 - \dfrac{\alpha}{2}; n_1 - 1, n_2 - 1\right)$

lub

$F_{\text{emp}} > F\left(\dfrac{\alpha}{2}; n_1 - 1, n_2 - 1\right)$

to hipotezę $H_0 : \sigma_1^2 = \sigma_2^2$ odrzucamy

Uwaga

$$F\left(1-\alpha;u,v\right)=\frac{1}{F\left(\alpha;v,u\right)}$$

Reguła: większa wariancja do licznika.

Jeżeli $S_1^2 > S_2^2$, to wyznaczana jest statystyka

$$F_{\text{emp}}=\frac{S_1^2}{S_2^2}$$

i hipoteza jest odrzucana, gdy

$$F_{\text{emp}}>F\left(\frac{\alpha}{2};n_1-1,n_2-1\right)$$

Jeżeli zaś $S_1^2 < S_2^2$, to wyznaczana jest statystyka

$$F_{\text{emp}}=\frac{S_2^2}{S_1^2}$$

i hipoteza jest odrzucana, gdy

$$F_{\text{emp}}>F\left(\frac{\alpha}{2};n_2-1,n_1-1\right)$$

Przykład. Dla sprawdzenia stabilności pracy maszyny pobrano dwie próbki: pierwszą w początkowym okresie eksploatacji oraz drugą po miesięcznym okresie pracy tej maszyny. Wykonano pomiary wylosowanych produktów i otrzymano wyniki: $n_1 = 25$, $\bar{x}_1 = 3.24$, $s_1^2 = 0.1447$ oraz $n_2 = 19$, $\bar{x}_2 = 3.19$, $s_2^2 = 0.1521$. Zbadać na tej podstawie czy maszyna nie rozregulowała się w trakcie pracy.

Populacja 1
produkcja maszyny w początkowym okresie

Populacja 2
produkcja maszyny po miesiącu eksploatacji

Cecha X
pomiar produktu

Założenia
cecha X ma w populacji 1 rozkład $N(\mu_1, \sigma_1^2)$
cecha X ma w populacji 2 rozkład $N(\mu_2, \sigma_2^2)$

Formalizacja
Stabilność pracy maszyny może być mierzona podobieństwem wytwarzanych produktów: im własności produktów są do siebie bardziej zbliżone, tym bardziej stabilna jest praca maszyny. Podobieństwo takie jest wyrażane wariancją cechy. Zatem stabilność pracy można wyrazić liczbowo jako wariancję interesującej cechy produktu, a problem stabilności jako zagadnienie weryfikacji hipotezy $H_0 : \sigma_1^2 = \sigma_2^2$

Technika statystyczna
Test F (poziom istotności $\alpha = 0.10$)

Obliczenia

$$F_{\text{emp}} = \frac{s_2^2}{s_1^2} = 1.051$$

Wartość krytyczna $F\,(0.05; 19, 24) = 2.114$

Odpowiedź: hipotezy nie odrzucamy

Wniosek: można uznać że maszyna nie rozregulowała się w trakcie pracy

$$H_0 : \sigma_1^2 \leq \sigma_2^2$$

Test F (poziom istotności α)

Statystyka testowa

$$F_{\text{emp}} = \frac{S_1^2}{S_2^2}$$

Wartość krytyczna $F(\alpha; n_1 - 1, n_2 - 1)$

Jeżeli $F_{\text{emp}} > F(\alpha; n_1 - 1, n_2 - 1)$
to hipotezę $H_0 : \sigma_1^2 \leq \sigma_2^2$ odrzucamy

Uwaga
W tym przypadku zasada „większa wariancja do licznika" nie ma sensu.

Porównanie dwóch rozkładów dwupunktowych

Założenia:
1. $X_1 \sim D(p_1), X_2 \sim D(p_2)$
2. X_1, X_2 są niezależne

$$H_0 : p_1 = p_2$$

Test przybliżony (poziom istotności α)

$$\hat{p}_1 = \frac{k_1}{n_1}, \quad \hat{p}_2 = \frac{k_2}{n_2}, \quad \hat{p} = \frac{(k_1 + k_2)}{(n_1 + n_2)}$$

Statystyka testowa

$$u_{\text{emp}} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}$$

$|u_{\text{emp}}| \geq u_{1-\alpha/2} \Longrightarrow H_0 : p_1 = p_2$ odrzucamy

Przykład. Celem badania było porównanie przygotowania z matematyki kandydatów na studia będących absolwentami liceów oraz techników. W tym celu spośród kandydatów zdających matematykę wylosowano 400 absolwentów liceów oraz 600 absolwentów techników. W wylosowanej grupie stwierdzono, że 385 absolwentów liceów oraz 501 absolwentów techników rozwiązało test wstępny. Czy można na tej podstawie sądzić, że przygotowanie w obu grupach absolwentów jest jednakowe?

Populacja 1:
absolwenci liceów zdający egzamin wstępny
Populacja 2:
absolwenci techników zdający egzamin wstępny

Cecha X: umiejętność rozwiązania testu (tak/nie)

Założenia:
cecha X ma w populacji 1 rozkład $D(p_1)$
cecha X ma w populacji 2 rozkład $D(p_2)$

Formalizacja
Weryfikacja hipotezy $H_0 : p_1 = p_2$

Technika statystyczna
Test przybliżony (poziom istotności $\alpha = 0.05$)

Obliczenia

$$n_1 = 400 \quad k_1 = 385 \quad \hat{p}_1 = 385/400 = 0.9625$$

$$n_2 = 600 \quad k_2 = 501 \quad \hat{p}_2 = 501/600 = 0.8350$$

$$\hat{p} = (385 + 501)/(400 + 600) = 0.886$$

$$u_{\text{emp}} = \frac{0.9625 - 0.8350}{\sqrt{0.886(1 - 0.886)\left(\frac{1}{400} + \frac{1}{600}\right)}} = 6.215.$$

Wartość krytyczna $u_{0.975} = 1.96$

Odpowiedź: hipotezę $H_0 : p_1 = p_2$ odrzucamy

Wniosek:
przygotowanie absolwentów liceów i techników z matematyki nie jest takie same.

$$H_0 : p_1 \leq p_2$$

Test przybliżony (poziom istotności α)

$$\hat{p}_1 = \frac{k_1}{n_1}, \quad \hat{p}_2 = \frac{k_2}{n_2}, \quad \hat{p} = \frac{(k_1 + k_2)}{(n_1 + n_2)}$$

Statystyka testowa

$$u_{\mathrm{emp}} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}$$

$u_{\mathrm{emp}} \geq u_{1-\alpha} \Longrightarrow H_0 : p_1 \leq p_2$ odrzucamy

Porównanie wielu rozkładów normalnych

Założenia:
1. $X_i \sim N(\mu_i, \sigma_i^2), \quad i = 1, \ldots, k$
2. $X_1, \ldots, X_k$ są niezależne

Czy $\mu_1 = \cdots = \mu_k$?

Czy $\sigma_1^2 = \cdots = \sigma_k^2$?

Próby: $X_{i1}, \ldots, X_{in_i}, \quad i = 1, \ldots, k$

$$\bar{X}_i, \quad \mathrm{var} X_i, \quad s_i^2 = \frac{\mathrm{var} X_i}{n_i - 1}; \quad i = 1, \ldots, k$$

$$H_0 : \mu_1 = \cdots = \mu_k$$

Założenie $\sigma_1^2 = \cdots = \sigma_k^2$

Test F (poziom istotności α)

Statystyka testowa

$$F_{\mathrm{emp}} = \frac{S_a^2}{S_e^2}$$

$$S_a^2 = \frac{1}{k-1} \sum_{i=1}^{k} n_i (\bar{X}_i - \bar{X})^2$$

$$S_e^2 = \frac{1}{N-k} \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$

$$\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}, \quad \bar{X} = \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{n_i} X_{ij}$$

$$N = \sum_{i=1}^{k} n_i$$

Jeżeli $F_{\text{emp}} > F\,(\alpha; k-1, N-k)$,
to hipotezę $H_0 : \mu_1 = \cdots = \mu_k$ odrzucamy.

Wniosek praktyczny:
przynajmniej jedna ze średnich $\mu_1, \ldots, \mu_k$ jest inna od pozostałych

Model analizy wariancji

$$X_{ij} = \mu_i + \varepsilon_{ij}$$

Błąd losowy $\varepsilon_{ij} \sim N(0, \sigma^2)$

Przykłady
<u>Plenność</u> kilku *odmian* pewnej rośliny uprawnej
<u>Wydajność</u> pracowników kilku *zakładów* pracy
<u>Zarobki</u> kilku *grup* społecznych

Czynnik: odmiana, zakład, grupa
Poziomy czynnika: badane odmiany, badane zakłady, badane grupy

Model analizy wariancji

$$X_{ij} = \mu + a_i + \varepsilon_{ij}$$

a_i — efekt i–tego poziomu czynnika: $\sum_{i=1}^k a_i = 0$

$$H_0 : a_1 = \cdots = a_k = 0, \qquad H_0 : \sum_{i=1}^k a_i^2 = 0$$

Tabela analizy wariancji

Źródło zmienności	Stopnie swobody	Sumy kwadratów	Średnie kwadraty	F_{emp}
Czynnik	$k-1$	var$A$	$S_a^2 = \frac{\text{var}A}{k-1}$	S_a^2/S_e^2
Błąd losowy	$N-k$	var$E$	$S_e^2 = \frac{varE}{N-k}$	
Ogółem	$N-1$	var$T$		

$$\text{var}A = \sum_{i=1}^k n_i(\bar{X}_i - \bar{X})^2, \text{var}E = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2,$$

$$\text{var}T = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2,$$

$$\text{var}A + \text{var}E = \text{var}T$$

Grupy jednorodne — podzbiory średnich, które można uznać za takie same

Procedury porównań wielokrotnych — postępowanie statystyczne zmierzające do podzielenia zbioru średnich na grupy jednorodne

Procedury: Tukeya, Scheffégo, Bonfferroniego, Duncana, Newmana–Kuelsa i inne.

Ogólna idea procedur porównań wielokrotnych $(n_1 = \cdots = n_k)$

NIR — najmniejsza istotna różnica

Jeżeli $|\bar{X}_i - \bar{X}_j| < NIR$, to uznajemy, że $\mu_i = \mu_j$.

Jeżeli
$|\bar{X}_i - \bar{X}_j| < NIR$
$|\bar{X}_i - \bar{X}_l| < NIR$
$|\bar{X}_l - \bar{X}_j| < NIR$,
to uznajemy, że $\mu_i = \mu_j = \mu_l$.

Badając w ten sposób wszystkie pary średnich próbkowych otrzymujemy podział zbioru średnich na grupy jednorodne.

Procedura Tukeya

Założenie: $n_1 = \cdots = n_k = n$

$$NIR = t(\alpha; k, N-k) S_e \sqrt{\frac{1}{n}}$$

$t(\alpha; k, N-k)$ — wartość krytyczna studentyzowanego rozstępu

Przypadek nierównolicznych prób
Jedna z modyfikacji procedury Tukeya

$$NIR_{ij} = t(\alpha; k, N-k) S_e \sqrt{\frac{1}{2}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

Przykład. Przeprowadzić analizę porównawczą wyników punktowych klasówki w grupach studenckich.

Populacje
Możemy wyodrębnić dziesięć populacji indeksowanych numerami grup studenckich

Cecha X
Ilość punktów uzyskanych na klasówce

Założenia
cecha X ma w i–tej populacji rozkład $N(\mu_i, \sigma_i^2)$
$(i = 1, \ldots, 10)$
$\sigma_1^2 = \cdots = \sigma_{10}^2$

Formalizacja
weryfikacja hpotezy $H_0 : \mu_1 = \cdots = \mu_{10}$

Techniki statystyczna
• Jednoczynnikowa analiza wariancji
• Porównania szczegółowe
Poziom istotności 0.05

Obliczenia

i	n_i	$\sum x_i$	$\sum x_i^2$
1	30	18.230	11.375950
2	30	16.672	9.596790
3	30	14.292	7.087458
4	30	18.879	12.069655
5	30	18.200	11.355982
6	30	19.568	13.172884
7	30	16.522	9.420960
8	30	19.134	12.514874
9	30	18.548	11.945964
10	30	16.521	9.304785
	300	176.566	107.845302

i	n_i	$\bar{x}_i$	$n_i(\bar{x}_i - \bar{x})^2$	$\text{var}x_i$
1	30	0.607667	0.010960	0.298187
2	30	0.555733	0.032315	0.331604
3	30	0.476400	0.377351	0.278749
4	30	0.629300	0.049809	0.189100
5	30	0.606667	0.009843	0.314649
6	30	0.652267	0.121782	0.409330
7	30	0.550733	0.042911	0.321744
8	30	0.637800	0.072757	0.311209
9	30	0.618267	0.026486	0.478354
10	30	0.550700	0.042986	0.206670
	N=300	$\bar{x}$=0.588553	var$A$=0.787199	var$E$=3.139595

$$\text{var}T = 107.845302 - 176.566^2/300 = 3.926794$$

Tabela analizy wariancji

Źródło zmienności	Stopnie swobody	Sumy kwadratów	Średnie kwadraty	F_{emp}
Grupa	9	0.787199	0.087467	8.079
Błąd losowy	290	3.139595	0.010826	
Ogółem	299	3.926794		

Wartość krytyczna

$$F\,(0.05; 9, 290) = 1.912$$

Odpowiedź:
hipotezę $H_0 : \mu_1 = \cdots = \mu_{10}$ odrzucamy

Wniosek:
przynajmniej jedna grupa uzyskała inną średnią liczbę punktów niż pozostałe

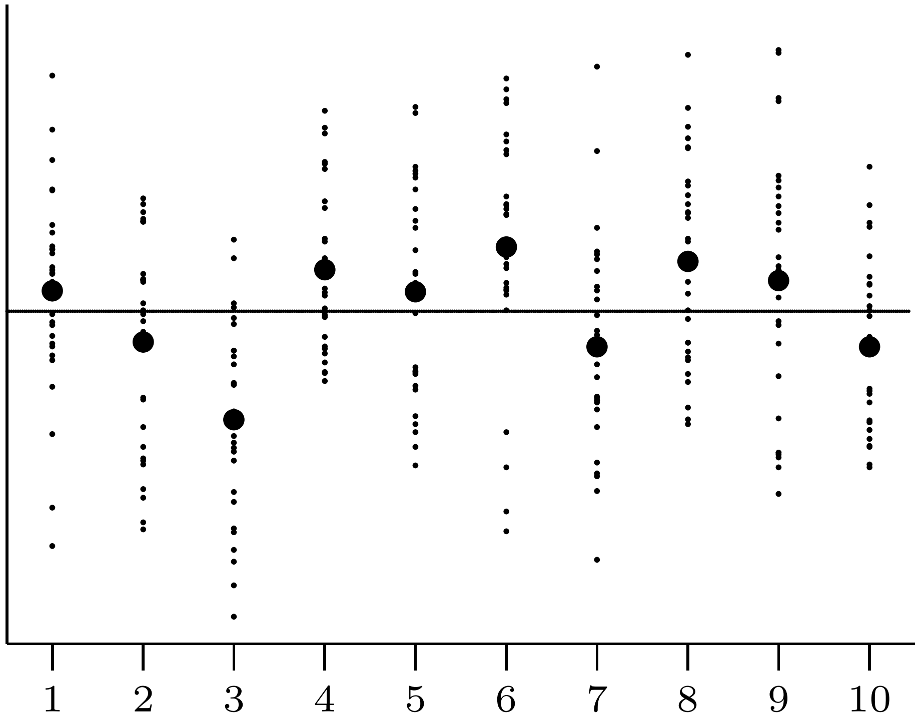
Wyznaczenie grup jednorodnych

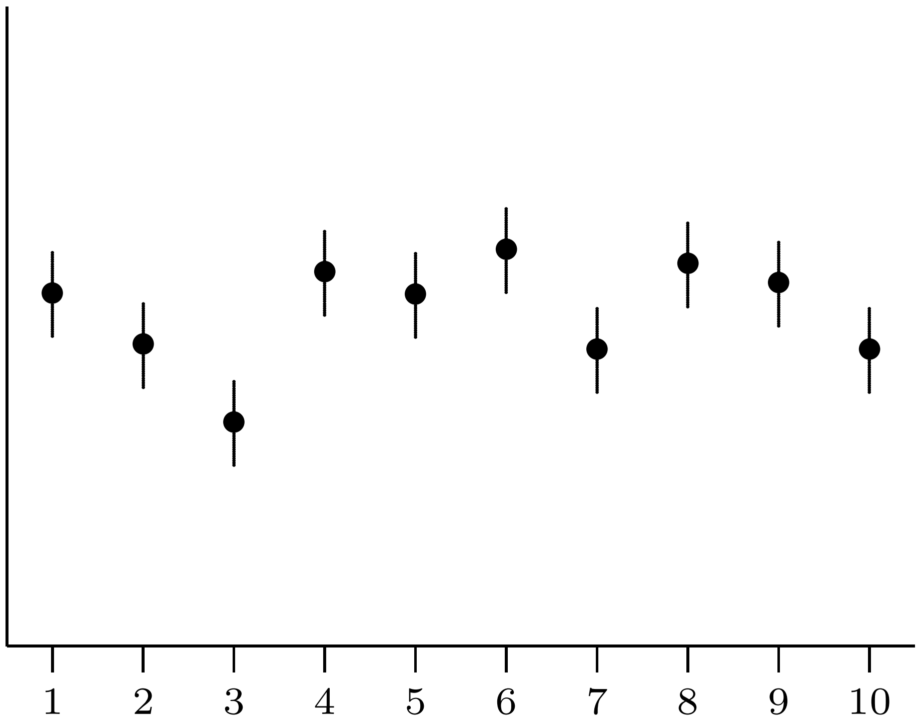
Procedura Tukeya ($\alpha = 0.05$)

Wartość krytyczna: $t(0.05; 10, 290) = 4.474$

$$NIR = 4.474 \cdot \sqrt{0.010826} \cdot \sqrt{\frac{1}{30}} = 0.084990$$

i	$\bar{x}_i$				
3	0.476400	*			
10	0.550700	*	*		
7	0.550733	*	*		
2	0.555733	*	*	*	
5	0.606667		*	*	*
1	0.607667		*	*	*
9	0.618267		*	*	*
4	0.629300		*	*	*
8	0.637800			*	*
6	0.652267				*


1
2
3
4
5
6
7
8
9
10


1
2
3
4
5
6
7
8
9
10

Porównanie wariancji
Cecha X_i ma rozkład normalny $N(\mu_i, \sigma_i^2)$
Średnie μ_i oraz wariancje σ_i^2 są nieznane

$$H_0 : \sigma_1^2 = \cdots = \sigma_k^2$$

Test Bartletta (poziom istotności α)
Statystyka testowa

$$M = (N-k)\ln\left(\frac{1}{N-k}\sum_{i=1}^{k}(n_i-1)S_i^2\right) - \sum_{i=1}^{k}(n_i-1)\ln S_i^2$$

$$S_i^2 = \frac{1}{n_i-1}\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)^2$$

Jeżeli $M > m(\alpha)$, to $H_0 : \sigma_1^2 = \cdots = \sigma_k^2$ odrzucamy.

$m(\alpha) =$

$$\frac{1}{c_1 - c}[(c_1 - c_3)m_1(\alpha; k, c_1) + (c_3 - c)m_2(\alpha; k, c_1)]$$

$$c_1 = \sum_{i=1}^{k} \frac{1}{n_i - 1} - \frac{1}{N - k}$$

$$c_3 = \sum_{i=1}^{k} \frac{1}{(n_i - 1)^3} - \frac{1}{(N - k)^3} \ ,$$

$$c = c_1^3/k^2 \ ,$$

$m_1(\alpha; k, c_1), \quad m_2(\alpha; k, c_1)$ są stablicowane
Jeżeli wszystkie $n_i > 4$, to statystyka testowa

$$\frac{M}{1 + c_1/(3(k-1))}$$

ma w przybliżeniu rozkład chi–kwadrat z $k-1$ stopniami swobody.
Jeżeli $c_1 = 0$, to

$$m_1(\alpha; k, c_1) = m_2(\alpha; k, c_1) = \chi^2(\alpha; k - 1)$$

Przypadek $n_1 = \cdots = n_k = n$

Test Cochrana (poziom istotności α)
Statystyka testowa

$$G = \frac{S^2_{\max}}{S^2_1 + \cdots + S^2_k}$$

$$S^2_{\max} = \max\{S^2_1, \ldots, S^2_k\}$$

Jeżeli $G > g(\alpha; k, n)$,
to $H_0 : \sigma^2_1 = \cdots = \sigma^2_k$ odrzucamy

Wartości krytyczne $g(\alpha; k, n)$ są podane w tablicach

Przypadek $n_1 = \cdots = n_k = n$

Test Hartleya (poziom istotności α)
Statystyka testowa

$$F_{\max} = \frac{S^2_{\max}}{S^2_{\min}}$$

$$S^2_{\min} = \min\{S^2_1, \ldots, S^2_k\}$$

Jeżeli $F_{\max} > f_{\max}(\alpha; k, n)$,
to $H_0 : \sigma^2_1 = \cdots = \sigma^2_k$ odrzucamy

Wartości krytyczne $f_{\max}(\alpha; k, n)$ są podane w tablicach

Przykład. W celu porównania zróżnicowania cen targowiskowych na jaja w czterech województwach w Polsce z każdego województwa wylosowano pewne ilości targowisk i zanotowano przeciętne ceny jaj na tych targowiskach. Po odpowiednich przeliczeniach uzyskano nastepujące wyniki

Województwo	Liczba targowisk n_i	Wariancja s_i^2
1	8	900
2	6	400
3	5	400
4	7	1600

Czy można na tej podstawie uznać, że zróżnicowanie cen w badanych województwach jest takie same?

Populacje
Są cztery populacje: targowiska w badanych województwach

Cecha X
przeciętna cena jaj na targowisku

Założenie
cecha w i–tej populacji ma rozkład $N(\mu_i, \sigma_i^2)$
$(i = 1, 2, 3, 4)$

Formalizacja

Miernikiem zróżnicowania cechy jest jej wariancja. Zatem problem analizy porównawczej zróżnicowania cen można zapisać jako zagadnienie weryfikacji hipotezy $H_0 : \sigma_1^2 = \cdots = \sigma_4^2$

Technika statystyczna

Test Bartletta (poziom istoności $\alpha = 0.05$)

Obliczenia

	n_i	$(n_i-1)s_i^2$	$(n_i-1)\ln s_i^2$	$1/(n_i-1)$	$1/(n_i-1)^3$
1	8	6300	47.6168	0.1429	0.0029
2	6	2000	29.9573	0.2000	0.0080
3	5	1600	23.9659	0.2500	0.0156
4	7	9600	44.2666	0.1667	0.0046
Razem	26	19500	145.8065	0.7595	0.0312

$$M = (26-4)\ln\left(\frac{19500}{26-4}\right) - 145.8065 = 3.5103$$

$$c_1 = 0.7595 - \frac{1}{(26-4)} = 0.7141$$

$$c_3 = 0.0312 - \frac{1}{(26-4)^3} = 0.0311$$

$$c = \frac{0.7141^3}{4^2} = 0.0228$$

Wartość krytyczna

$$m_1(0.05; 4, 0.7141) = 8.4630$$

$$m_2(0.05; 4, 0.7141) = 8.0972$$

$$m(0.05) = \frac{(0.7141-0.0311)8.4630+(0.0311-0.0228)8.0972}{0.7141-0.0228} = 8.4586$$

Odpowiedź: nie ma podstaw do odrzucenia weryfikowanej hipotezy

Wniosek: zróżnicowanie cen targowiskowych w badanych województwach można uznać za takie same.