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:	Wniosek o hipotezie H_0	
hipoteza 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

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

Wartość krytyczna $t(\alpha; n-1)$

Jeżeli $|t_{\mathrm{emp}}| > t(\alpha; n-1)$, 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

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

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

Odpowiedź: hipotezę odrzucamy

Wniosek: maszyna uległa rozregulowaniu

Moc testu

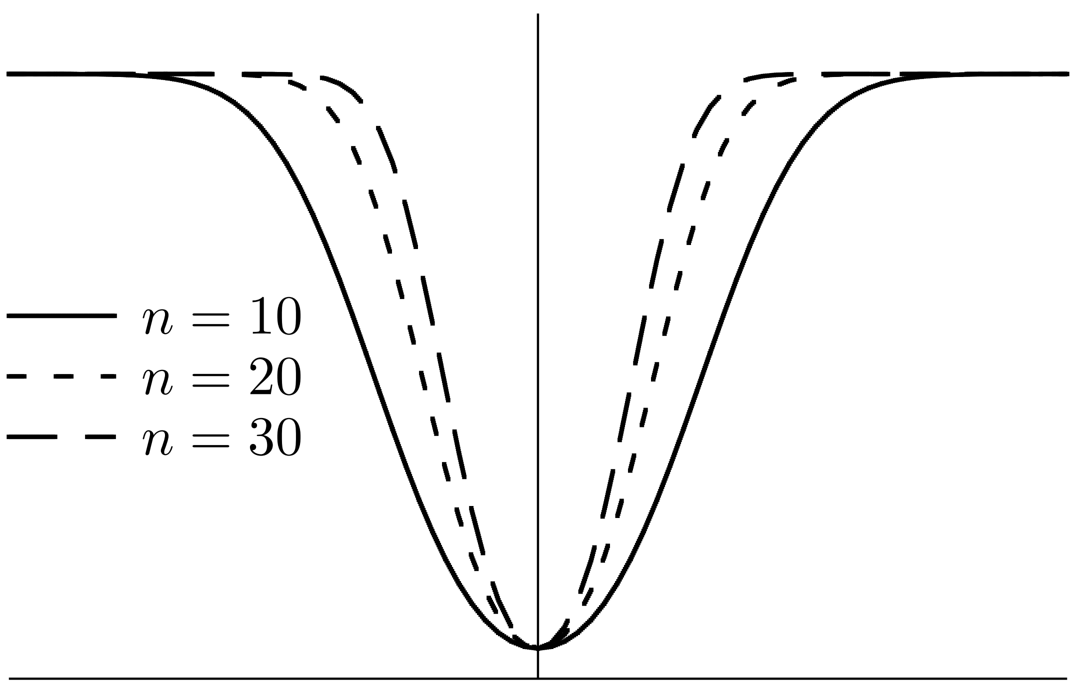
$$\text{Moc testu} = 1 - P\{\text{błąd II rodzaju}\}$$

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

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

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

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



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.

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

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

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

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

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

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

Wartość krytyczna $u_{1-\alpha/2}$

Jeżeli $|u_{\text{emp}}| > u_{1-\alpha/2}$, 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