

Testy oparte na ilorazie wiarygodności

Problem testowania hipotez (przypomnienie)

Niech $(\mathcal{X}, \mathcal{B}, \mathcal{P} = \{P_\theta : \theta \in \Theta\})$ będzie przestrzenią statystyczną przy czym $\Theta = \Theta_0 \cup \Theta_1$ i $\Theta_0 \cap \Theta_1 = \emptyset$.

Problem. Na podstawie próby (wyniku eksperymentu) $X \in \mathcal{X}$ zweryfikować hipotezę

$$H_0: \theta \in \Theta_0 \text{ wobec alternatywy } H_1: \theta \in \Theta_1.$$

Rozwiązaniem powyższego problemu jest funkcja $\varphi: \mathcal{X} \rightarrow [0, 1]$ zwana testem statystycznym (zrandomizowanym lub niezrandomizowanym).

Załóżmy ponadto, że rodzina rozkładów $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ jest dominowana przez pewną σ -skończoną miarę μ (zwykle miarę liczącą lub Lebesgue'a) i oznaczmy gęstość rozkładu P_θ względem miary μ przez $p_\theta = \frac{dP_\theta}{d\mu}$. Rodzinę $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ można więc utożsamić z rodziną gęstości $\mathcal{P} = \{p_\theta : \theta \in \Theta\}$.

Wiadomo, że

- Lemat Neymana-Pearsona podaje sposób konstrukcji testu najmocniejszego (czyli optymalnego w tym podejściu) hipotezy prostej przeciwko prostej alternatywie
- Czasami udaje się skonstruować test najmocniejszy hipotezy prostej przeciwko złożonej alternatywie- np., gdy statystyka testowa testu najmocniejszego hipotezy prostej przeciwko prostej alternatywie nie zależy od tej alternatywy dla alternatyw z pewnej klasy.
- Rzadziej, zwykle w jednowymiarowych rodzinach wykładniczych, wykorzystując ich pewne „monotoniczne” własności, można skonstruować test jednostajnie najmocniejszy hipotezy złożonej przeciwko złożonej alternatywie- przykłady na ćwiczeniach.
- Podobnie rzadko potrafimy konstruować testy optymalne w podejściu minimaksowym lub bayesowskim.

Pytanie. Czy istnieje jakaś heurystyczna metoda konstrukcji testów niekoniecznie optymalnych, ale praktycznie użytecznych (podobnie jak metoda NW w teorii estymacji)

Okazuje się, że w problemie testowania hipotez **konceptcja największej wiarygodności** jest również bardzo owocna.

Niech $L(\theta; X)$ będzie funkcją wiarygodności w przestrzeni $(\mathcal{X}, \mathcal{B}, \mathcal{P} = \{P_\theta : \theta \in \Theta\})$

Def. Wiarygodnością hipotezy $H_i: \theta \in \Theta_i$ gdy zaobserwowano $X \in \mathcal{X}$ nazywamy liczbę

$$L_{H_i}(X) = \sup_{\theta \in \Theta_i} L(\theta, X).$$

Uwaga. Wiarygodność hipotezy H_i możemy interpretować w przypadku dyskretnym jako największe prawdopodobieństwo, przy prawdziwości H_i , uzyskania zaobserwowanego $X \in \mathcal{X}$. Podobną interpretację (z oczywistymi modyfikacjami związanymi z interpretacją gęstości) mamy w przypadku ciągłym.

Jeżeli weryfikujemy hipotezę $H_0: \theta \in \Theta_0$ wobec alternatywy $H_1: \theta \in \Theta_1$, to należy zakwestionować hipotezę H_0 , gdy bardziej wiarygodna jest hipoteza H_1 , tzn. gdy

$$\frac{L_{H_1}(X)}{L_{H_0}(X)} > \lambda_0 \geq 1,$$

gdzie λ_0 jest tak dobraną stałą, aby rozmiar testu nie przekraczał poziomu istotności α , czyli

$$\sup_{\theta \in \Theta_0} P_{\theta} \left(X : \frac{L_{H_1}(X)}{L_{H_0}(X)} > \lambda_0 \right) \leq \alpha.$$

Niech $\dot{\theta}$ będzie taką wartością parametru $\theta \in \Theta_0$ (o ile istnieje), że $L(\dot{\theta}, X) = L_{H_0}(X) = \sup_{\theta \in \Theta_0} L(\theta, X)$,

czyli zakładamy, że w punkcie $\dot{\theta}$ realizuje się wiarygodność hipotezy zerowej (czyli jest osiągany odpowiedni kres) i niech $\hat{\theta}$ będzie estymatorem największej wiarygodności parametru θ czyli

$$L(\hat{\theta}, X) = \sup_{\theta \in \Theta} L(\theta, X).$$

Estymator $\hat{\theta}$ nazywamy bezwarunkowym ENW a $\dot{\theta}$ nazywamy warunkowym (przy prawdziwości H_0) ENW parametru θ .

Widać, że dla $\lambda_0 \geq 1$ mamy następującą równoważność

$$\frac{L_{H_1}(X)}{L_{H_0}(X)} > \lambda_0 \Leftrightarrow \frac{L(\hat{\theta}, X)}{L(\dot{\theta}, X)} > \lambda_0$$

bo $L_{H_1}(X) > L_{H_0}(X) \Leftrightarrow \sup_{\theta \in \Theta} L(\theta, X) = \sup_{\theta \in \Theta_1} L(\theta, X)$

Def. Wielkość $\lambda(X) = \frac{L(\hat{\theta}, X)}{L(\dot{\theta}, X)} = \frac{\sup_{\theta \in \Theta} L(\theta, X)}{\sup_{\theta \in \Theta_0} L(\theta, X)}$ nazywamy **ilorazem wiarygodności** a test H_0

przeciwko H_1 o obszarze krytycznym (obszarze odrzucenia H_0) postaci

$$C = \{X: \lambda(X) > \lambda_0\}$$

gdzie $\lambda_0 \geq 1$ jest odpowiednio dobraną stałą nazywamy **testem opartym na ilorazie wiarygodności**.

Przykład 1. Test t -Studenta jako test oparty na ilorazie wiarygodności.

Niech $X=(X_1, \dots, X_n)$ będzie próbą prostą z rozkładu $N(m, \sigma^2)$ z nieznanymi parametrami m i σ^2 .

Testujemy hipotezę

$$H_0: m=m_0 \text{ wobec alternatywy } H_1: m \neq m_0.$$

Funkcja wiarygodności ma postać

$$L(m, \sigma; X) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - m)^2}.$$

Stąd $(\hat{m}, \hat{\sigma}) = (\bar{X}, \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}) = ENW[(m, \sigma)]$. Wobec tego

$$L(\hat{m}, \hat{\sigma}; X) = (2\pi\hat{\sigma}^2)^{-\frac{n}{2}} e^{-\frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n (X_i - \bar{X})^2} = (2\pi\hat{\sigma}^2)^{-\frac{n}{2}} e^{-\frac{n}{2}}.$$

Podobnie przy prawdziwości hipotezy H_0 funkcja wiarygodności ma postać

$$L(\sigma; X) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - m_0)^2}. \text{ Skąd łatwo widać, że } \hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - m_0)^2} = ENW[\sigma] \text{ i}$$

$$L(\hat{\sigma}; X) = (2\pi\hat{\sigma}^2)^{-\frac{n}{2}} e^{-\frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n (X_i - m_0)^2} = (2\pi\hat{\sigma}^2)^{-\frac{n}{2}} e^{-\frac{n}{2}}.$$

Wobec tego

$$\begin{aligned} C = \{X : \lambda(X) > k_1\} &= \{X : \frac{\sum_{i=1}^n (X_i - m_0)^2}{\sum_{i=1}^n (X_i - \bar{X})^2} > k_2\} = \{X : \frac{\sum_{i=1}^n ((X_i - \bar{X}) + (\bar{X} - m_0))^2}{\sum_{i=1}^n (X_i - \bar{X})^2} > k_2\} = \\ &= \{X : \frac{\sum_{i=1}^n (X_i - \bar{X})^2 + n(\bar{X} - m_0)^2}{\sum_{i=1}^n (X_i - \bar{X})^2} > k_2\} = \{X : 1 + \frac{n(\bar{X} - m_0)^2}{\sum_{i=1}^n (X_i - \bar{X})^2} > k_2\} = \{X : \frac{n(\bar{X} - m_0)^2}{\sum_{i=1}^n (X_i - \bar{X})^2} > k_3\} = \\ &= \{X : |t(X)| = \frac{\sqrt{n} |\bar{X} - m_0|}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}} > k\} \end{aligned}$$

Wiadomo, że przy prawdziwości H_0 statystyka testowa $t(X) = \frac{\sqrt{n}(\bar{X} - m_0)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}}$ ma rozkład t_{n-1}

Studenta o $n-1$ stopniach swobody, gdyż zmienne losowe $\frac{\sqrt{n}(\bar{X} - m_0)}{\sigma}$ i $\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2}$ mają

odpowiednio rozkłady $N(0,1)$ i χ_{n-1}^2 oraz są one niezależne (tw Fishera -wniosek z tw. Basu).

Wobec tego test oparty na ilorazie wiarygodności hipotezy $H_0: m=m_0$ wobec alternatywy $H_1: m \neq m_0$

ma obszar krytyczny postaci $C = \{X : |t(X)| = \frac{\sqrt{n} |\bar{X} - m_0|}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}} > t(1 - \frac{\alpha}{2}, n-1)\}$ gdzie $t(p, n)$ jest

kwantylem rzędu p rozkładu t -Studenta o n stopniach swobody.

Przykład 2 Testowanie równości średnich w rozkładach normalnych.

Niech $X=(X_1, \dots, X_m)$ i $Y=(Y_1, \dots, Y_n)$ będą niezależnymi próbami prostymi z rozkładów $N(m_x, \sigma^2)$ i $N(m_y, \sigma^2)$ z nieznanymi parametrami m_x , m_y i σ^2 . Testujemy hipotezę

$H_0: m_x = m_y = \mu$ wobec alternatywy $H_1: m_x \neq m_y$

Funkcja wiarygodności ma postać

$$L(m_x, m_y, \sigma; \mathbf{X}, \mathbf{Y}) = (2\pi\sigma^2)^{-\frac{m+n}{2}} e^{-\frac{1}{2\sigma^2} \{ \sum_{i=1}^m (X_i - m_x)^2 + \sum_{j=1}^n (Y_j - m_y)^2 \}}.$$

a jej logarytm

$$l(m_x, m_y, \sigma; \mathbf{X}, \mathbf{Y}) = -\frac{m+n}{2} \ln(2\pi) - (m+n) \ln \sigma - \frac{1}{2\sigma^2} \{ \sum_{i=1}^m (X_i - m_x)^2 + \sum_{j=1}^n (Y_j - m_y)^2 \}$$

Łatwo widać, że

$$(\hat{m}_x, \hat{m}_y, \hat{\sigma}) = (\bar{X}, \bar{Y}, \sqrt{\frac{1}{m+n} \{ \sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2 \}}) = ENW[(m_x, m_y, \sigma)].$$

Stąd $L(\hat{m}_x, \hat{m}_y, \hat{\sigma}; \mathbf{X}, \mathbf{Y}) = (2\pi\hat{\sigma}^2)^{-\frac{m+n}{2}} e^{-\frac{m+n}{2}}$.

W przypadku prawdziwości $H_0: m_x = m_y = \mu$ funkcja wiarygodności przybiera postać

$$L(\mu, \sigma; \mathbf{X}, \mathbf{Y}) = (2\pi\sigma^2)^{-\frac{m+n}{2}} e^{-\frac{1}{2\sigma^2} \{ \sum_{i=1}^m (X_i - \mu)^2 + \sum_{j=1}^n (Y_j - \mu)^2 \}}.$$

Stąd

$$l(\mu, \sigma; \mathbf{X}, \mathbf{Y}) = -\frac{m+n}{2} \ln(2\pi) - (m+n) \ln \sigma - \frac{1}{2\sigma^2} \{ \sum_{i=1}^m (X_i - \mu)^2 + \sum_{j=1}^n (Y_j - \mu)^2 \}$$

i w konsekwencji

$$(\dot{\mu}, \dot{\sigma}) = (\frac{m}{m+n} \bar{X} + \frac{n}{m+n} \bar{Y}, \sqrt{\frac{1}{m+n} \{ \sum_{i=1}^m (X_i - \dot{\mu})^2 + \sum_{j=1}^n (Y_j - \dot{\mu})^2 \}}) = ENW[(\mu, \sigma)].$$

Wobec tego

$$L(\dot{\mu}, \dot{\sigma}; \mathbf{X}, \mathbf{Y}) = (2\pi\dot{\sigma}^2)^{-\frac{m+n}{2}} e^{-\frac{m+n}{2}}$$

Test oparty na ilorazie wiarygodności H_0 przeciwko H_1 ma obszar krytyczny postaci

$$\begin{aligned}
C = \{(\mathbf{X}, \mathbf{Y}) : \lambda(\mathbf{X}, \mathbf{Y}) > k_1\} &= \{(\mathbf{X}, \mathbf{Y}) : \frac{\sum_{i=1}^m (X_i - \hat{\mu})^2 + \sum_{j=1}^n (Y_j - \hat{\mu})^2}{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2} > k_2\} = \\
&= \{(\mathbf{X}, \mathbf{Y}) : \frac{\sum_{i=1}^m ((X_i - \bar{X}) + (\bar{X} - \hat{\mu}))^2 + \sum_{j=1}^n ((Y_j - \bar{Y}) + (\bar{Y} - \hat{\mu}))^2}{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2} > k_2\} = \\
&= \{(\mathbf{X}, \mathbf{Y}) : 1 + \frac{m(\bar{X} - \hat{\mu})^2 + n(\bar{Y} - \hat{\mu})^2}{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2} > k_2\} = \{(\mathbf{X}, \mathbf{Y}) : \frac{\frac{mn}{m+n}(\bar{X} - \bar{Y})^2}{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2} > k_3\} = \\
&= \{(\mathbf{X}, \mathbf{Y}) : |t(\mathbf{X}, \mathbf{Y})| = \frac{\sqrt{\frac{mn(n+m-2)}{m+n}} |\bar{X} - \bar{Y}|}{\sqrt{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2}} > k\}
\end{aligned}$$

Wiadomo, że przy prawdziwości hipotezy H_0 statystyka $t(\mathbf{X}, \mathbf{Y})$ ma rozkład t -Studenta t_{n+m-2} o $n+m-2$ stopniach swobody, gdyż:

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{m}), \quad \bar{Y} \sim N(\mu, \frac{\sigma^2}{n}), \quad \frac{1}{\sigma^2} \sum_{i=1}^m (X_i - \bar{X})^2 \sim \chi_{m-1}^2, \quad \frac{1}{\sigma^2} \sum_{j=1}^n (Y_j - \bar{Y})^2 \sim \chi_{n-1}^2 \text{ i wszystkie zmienne}$$

są niezależne. Wobec tego $\bar{X} - \bar{Y} \sim N(0, (\frac{1}{m} + \frac{1}{n})\sigma^2) = N(0, \frac{m+n}{mn}\sigma^2)$

$$\frac{1}{\sigma^2} \sum_{i=1}^m (X_i - \bar{X})^2 + \frac{1}{\sigma^2} \sum_{j=1}^n (Y_j - \bar{Y})^2 \sim \chi_{m+n-2}^2 \text{ (tw. o dodawaniu)}$$

i obie zmienne są niezależne.

$$\text{Stąd z definicji rozkładu } t\text{-Studenta } t(X, Y) = \frac{\sqrt{\frac{mn(n+m-2)}{m+n}} (\bar{X} - \bar{Y})}{\sqrt{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2}} \sim t_{m+n-2}.$$

Graniczny rozkład ilorazu wiarygodności

Tw. Przy pewnych warunkach regularności (regularny model Cramera-Rao) statystyka $2\ln\lambda(X)$ ma dla każdego $\theta \in \Theta_0$ (czyli przy prawdziwości H_0) graniczny rozkład χ_r^2 o r stopniach swobody, gdzie r jest liczbą ograniczeń nałożonych na θ potrzebnych do wyspecyfikowania H_0 .

Uwaga. W przykładach 1 i 2 statystyka $2\ln\lambda(X)$ ma graniczny rozkład χ_1^2

Uwaga. Aproksymacja $\sqrt{2\chi_n^2} - \sqrt{2n-1} \sim N(0,1)$ (dla dostatecznie dużych n)

Wprowadzenie do ANOVA

Wiemy już jak można testować hipotezę

$$H_0: m_x = m_y$$

o równości średnich wobec alternatywy

$$H_1: m_x \neq m_y$$

Problem. Jak testować hipotezę o równości średnich w k populacjach **zależnych lub niezależnych**

Dla ustalenia uwagi rozważymy przypadek prób niezależnych. Dokładniej; na podstawie k **niezależnych** prób prostych z odpowiednich rozkładów normalnych

$$X_{11}, \dots, X_{1n_1} \text{ iid } N(m_1, \sigma^2)$$

$$X_{21}, \dots, X_{2n_2} \text{ iid } N(m_2, \sigma^2)$$

.....

$$X_{k1}, \dots, X_{kn_k} \text{ iid } N(m_k, \sigma^2)$$

z nieznaną wariancją σ^2 zweryfikować hipotezę

$$H_0: m_1 = m_2 = \dots = m_k$$

wobec alternatywy

$$H_1: \sim H_0 \text{ (inaczej } \exists (i,j) \ m_i \neq m_j \text{)}.$$

Ponieważ H_0 jest koniunkcją hipotez $H_0^{ij}: m_i = m_j \ 1 \leq i < j \leq k$ problem sprowadza się do jednoczesnego testowania $\binom{k}{2}$ hipotez $H_0^{ij}: m_i = m_j$, które potrafimy już testować.

Wydawać by się mogło, że wystarczy $\binom{k}{2}$ **razy zastosować test o równości 2 średnich** w rozkładach normalnych i odrzucić H_0 , gdy są podstawy do odrzucenia przynajmniej jednej z hipotez H_0^{ij} , $1 \leq i < j \leq k$. Okazuje się, że **taka procedura testowa jest nieefektywna i pojawiają się problemy z poziomem istotności złożonego testu.**

Rozważmy nieco dokładniej problem jednoczesnego testowania m hipotez. Przyjmijmy oznaczenia

$(\mathcal{X}, \mathcal{B}, \mathcal{P} = \{P_\theta : \theta \in \Theta\})$ - przestrzeń statystyczna

$H_0^i: \theta \in \Theta_0^i$ - i -ta szczegółowa hipoteza zerowa

$H_1^i: \theta \notin \Theta_0^i$ - i -ta szczegółowa hipoteza alternatywna

φ_i - test H_0^i przeciwko H_1^i o obszarze krytycznym C_i i obszarze akceptacji $A_i = \mathcal{X} - C_i$ na poziomie istotności α_i . Przyjmijmy, że $\forall 1 \leq i \leq m \ \alpha_i = \alpha$.

Testujemy

$$H_0 : \forall 1 \leq i \leq m \quad \theta \in \Theta_0^i \quad \text{wobec alternatywy}$$

$$H_1 : \exists 1 \leq i \leq m \quad \theta \notin \Theta_0^i \quad \text{będącej zaprzeczeniem } H_0$$

Inaczej testujemy $H_0 : \theta \in \bigcap_{i=1}^m \Theta_0^i$ wobec alternatywy $H_1 : \theta \notin \bigcap_{i=1}^m \Theta_0^i$

Naturalny test φ hipotezy H_0 ma obszar krytyczny $C = \bigcup_{i=1}^m C_i$ i obszar akceptacji $A = \bigcap_{i=1}^m A_i$ - (inaczej

$$\varphi(X) = \mathbf{1}_{\bigcup_{i=1}^m C_i}(X).$$

Wobec tego mamy $P_{H_0}(C) = P_{H_0}\left(\bigcup_{i=1}^m C_i\right) \leq \sum_{i=1}^m P_{H_0}(C_i) = m\alpha$.

Rozmiar testu złożonego φ jest więc zdecydowanie większy niż rozmiary testów szczegółowych. Wykonując 20 porównań na poziomie $\alpha=0.05$ w przypadku równości wszystkich średnich i tak przeciętnie $20 \cdot 0.05=1$ raz zaobserwujemy „falszywie istotną różnicę”. **Uwagi o FDR Benjaminiego-Hochberga.**

Wniosek. Potrzebny jest inny (specjalny) test do jednoczesnego testowania hipotez

Intuicja. Oznaczając przez $\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$ próbkową średnią w i -tej grupie która jest „dobrym” estymatorem m_i

$i=1, \dots, k$ spodziewamy się, że w przypadku prawdziwości $H_0: m_1=m_2=\dots=m_k$ średnie te nie powinny się wiele różnić. Wydaje się rozsądne przyjąć jako statystykę testową jakąś miarę zróżnicowania tych średnich na

przykład $T_1 = \sum_{i=1}^k (\bar{X}_i - \bar{X})^2$, gdzie $\bar{X} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} X_{ij}$ jest średnią próbkową wszystkich obserwacji. Statystyką

testową mogła by też być jedna ze statystyk $T_2 = \sum_{i=1}^k |\bar{X}_i - \bar{X}|$, $T_3 = \sum_{i=1}^k |\bar{X}_i - \text{Med}\{\bar{X}_1, \dots, \bar{X}_k\}|$ lub

$T_4 = \max\{X_1, \dots, X_k\} - \min\{X_1, \dots, X_k\}$. Za wyborem T_1 przemawia jej postać formy kwadratowej, która może mieć jakiś rozkład χ^2 .

Konstrukcja testu równości k średnich w rozkładach normalnych opartego na ilorazie wiarygodności.

(można pominąć, gdyż jest to szczególny przypadek testowania tzw. hipotezy liniowej która będzie omówiona w kolejnym wykładzie)

Oznaczając przez $\mathbf{X}$ wszystkie obserwacje w próbie mamy funkcję wiarygodności postaci

$$L(m_1, \dots, m_k, \sigma^2; \mathbf{X}) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - m_i)^2} \quad \text{skąd łatwo otrzymujemy estymatory NW}$$

$$\hat{m}_i = \bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}, i=1, \dots, k, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 \quad \text{ i } L(\hat{m}_1, \dots, \hat{m}_k, \hat{\sigma}^2; \mathbf{X}) = (2\pi\hat{\sigma}^2)^{-\frac{n}{2}} e^{-\frac{n}{2}}.$$

Podobnie w przypadku prawdziwości H_0 oznaczając przez m wspólną średnią mamy funkcję wiarygodności

$$L(m, \sigma^2; \mathbf{X}) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - m)^2}, \quad \text{ dla której znajdujemy argument który ją maksymalizuje}$$

$$\hat{m} = \bar{X} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} X_{ij} = \sum_{i=1}^k \frac{n_i}{n} \bar{X}_i \quad \text{ i } \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2. \quad \text{ Stąd } L(\hat{m}, \hat{\sigma}^2; \mathbf{X}) = (2\pi\hat{\sigma}^2)^{-\frac{n}{2}} e^{-\frac{n}{2}}.$$

Iloraz wiarygodności jest więc postaci $\lambda(\mathbf{X}) = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}^2} \right)^{\frac{n}{2}}$. Stąd test rozważanej H_0 ma następujący zbiór

krytyczny:

$$C = \{X : \lambda(\mathbf{X}) = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}^2} \right)^{\frac{n}{2}} > k_1\} = \{X : \frac{\hat{\sigma}^2}{\hat{\sigma}^2} > k_2\} = \{X : \frac{\hat{\sigma}^2 - \hat{\sigma}^2}{\hat{\sigma}^2} > k_3\}$$

Zauważmy, że

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

Nazwijmy odpowiednie sumy

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2 \quad \text{całkowita suma kwadratów (SS}_{\text{total}})$$

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 \quad \text{suma kwadratów wewnątrz grup (SS}_{\text{within}})$$

$$\sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2 \quad \text{sumę kwadratów pomiędzy grupami (SS}_{\text{between}})$$

Jeśli podzielimy każdą z sum przez n to otrzymamy odpowiednio próbkowe wariancje

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2 \quad \text{próbkowa wariancja całkowita}$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 \quad \text{próbkowa wariancja wewnątrz grup}$$

$$\frac{1}{n} \sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2 \quad \text{próbkowa wariancja pomiędzy grupami}$$

Uzyskaliśmy dekompozycję, którą nazywamy analizą wariancji (ANOVA)

Wariancja całkowita	=	Wariancja wewnątrz grup	+	Wariancja pomiędzy grupami
---------------------	---	-------------------------	---	----------------------------

Uwaga. Test hipotezy H_0 o równości średnich oparty na ilorazie wiarygodności każe odrzucić

H_0 , jeśli próbkowa wariancja pomiędzy grupami jest duża na tle próbkowej wariancji wewnątrz grup

Konstrukcja testu ANOVA równości średnich jest szczególnym przypadkiem konstrukcji testu dla ogólnej hipotezy liniowej, dla której pokazane będą wszystkie szczegóły.