MNOŽICE ZAUPANJA, NAPOVEDNE MNOŽICE

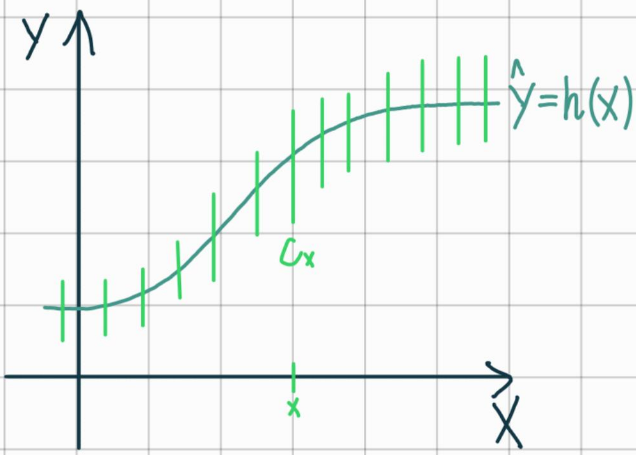
Spet: Opazimo X, zanima nas Y.

Želimo sklepati $\underset{h_{min}(X)}{\underset{\shortparallel}{Y_{min}}} < Y < \underset{h_{max}(X)}{\underset{\shortparallel}{Y_{max}}}$.

Če je $Y = y = g(\theta)$ deterministična količina, temu pravimo interval zaupanja, sicer pa je to napovedni model.

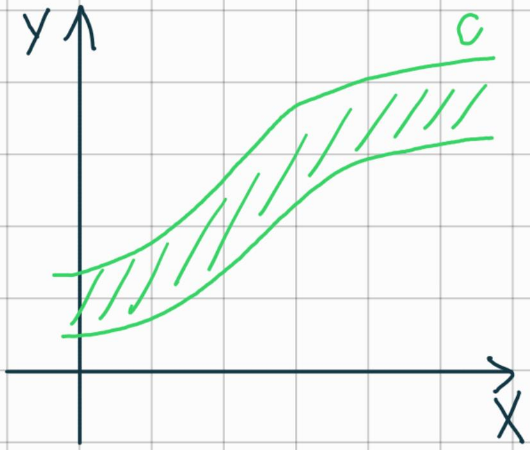
$$\forall \theta \in \Theta: \; P_\theta(Y_{min} < Y < Y_{max}) \geq 1-\alpha$$

Splošneje lahko sklepamo, da je $Y \in C_x$, kjer je C_x, $x \in \dots$, nabor nekih množic.



Ideje za konstrukcijo:

1) Celovit pogled:



$$C = \{(x,y)\ ;\ y \in C_x\}$$

$$P_\theta((X,Y) \in C) \underset{\approx}{\underset{=}{\geq}} 1-\alpha$$

2) Mera razhajanja:

$\rho(x,y)$ ... mera razhajanja

$$C = \{(x,y)\ ;\ \rho(x,y) < c_\alpha\}$$

Če imamo srečo, če ρ pivotna funkcija, kar pomeni, da se porazdelitev slučajne spremenljivke $\rho(X,Y)$ ne spreminja s θ.

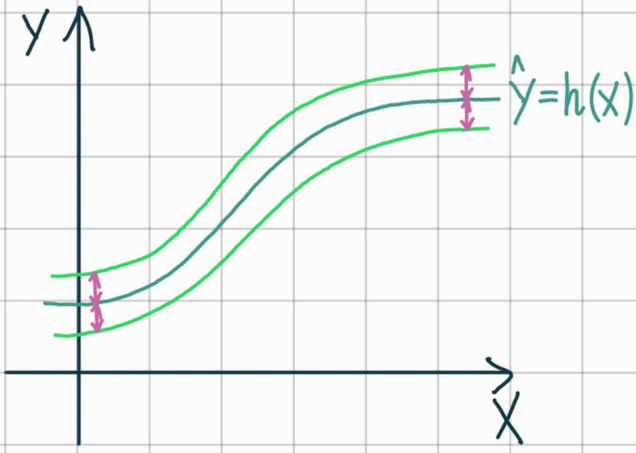
Tedaj lahko c_α dobimo kot minimalni c, za katerega je $P_{\boxed{\theta}}(\rho(X,Y) \leq c) \geq 1-\alpha$.

Tedaj je tudi $P_\theta(\rho(X,Y) \leq c_\alpha) \geq 1-\alpha$, često pa tudi $P_\theta(\rho(X,Y) < c_\alpha) \geq 1-\alpha$.

3) Uporaba točkovnih napovedi/cenilk:

$$\rho(x,y) := |h(x)-y|$$

Kjer je $\hat{Y} = h(X)$ napoved za Y.



Splošneje lahko vzamemo $\rho(x,y) = d(h(x), y)$.

Ponovimo standardni primer:

Opazimo $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$ neodvisne.

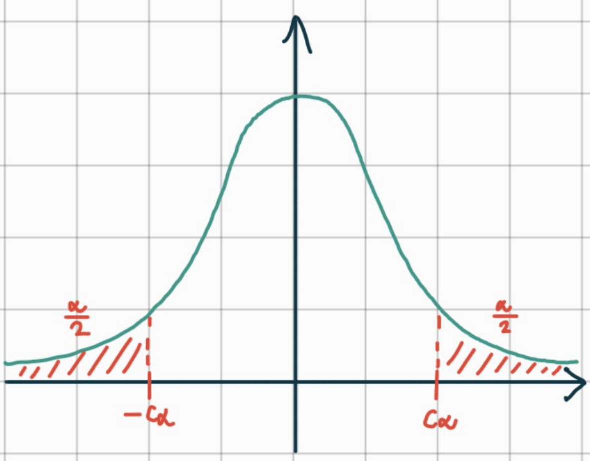
Recimo, da poznamo σ in ocenjujemo μ.

Za oceno $\hat{\mu} = \overline{X}$ dobimo:

$\rho((x_1, x_2, \ldots, x_n), \mu) = \left|\frac{x_1+x_2+\ldots+x_n}{n} - \mu\right|$

To je pivotna funkcija, ker je $\rho((X_1, X_2, \ldots, X_n), \mu) = |\overline{X} - \mu|$ in $\overline{X} - \mu \sim N(0, \frac{\sigma^2}{n})$, ne glede na $\mu = \theta$.

$c_\alpha := \Phi^{-1}(1 - \frac{\alpha}{2}) \frac{\sigma}{\sqrt{n}}$



$\overline{X} - c_\alpha < \mu < \overline{X} + c_\alpha$

Napovedni interval za $X_{n+1} \sim N(\mu, \sigma^2)$, neodvisne od $X_1, X_2, \ldots, X_n$:

$$\left.\begin{aligned} \overline{X} - \mu &\sim N(0, \tfrac{\sigma^2}{n}) \\ X_{n+1} - \mu &\sim N(0, \sigma^2) \end{aligned}\right\} \text{neodvisni}$$

$$\Rightarrow X_{n+1} - \overline{X} \sim N\left(0, \left(1+\frac{1}{n}\right)\sigma^2\right)$$

Torej je napovedni interval:

$$\overline{X} - c_\alpha^{nap} < X_{n+1} < \overline{X} + c_\alpha^{nap}$$

$$c_\alpha^{nap} = \phi^{-1}\left(1-\frac{\alpha}{2}\right)\sigma\sqrt{1+\frac{1}{n}}$$

Kaj pa, če σ ne poznamo?

Iz primera 7.14 vemo, da je $\frac{\overline{X}-\mu}{\hat{\sigma}_+}\sqrt{n} \sim Student(n-1)$, ne glede na μ in σ.

Interval zaupanja bo zdaj spet $\overline{X} - c_\alpha < \mu < \overline{X} + c_\alpha$, kjer pa bo:

$$c_\alpha := F^{-1}_{Student(n-1)}\left(1-\frac{\alpha}{2}\right)\frac{\hat{\sigma}_+}{\sqrt{n}}$$

Podobno dobimo tudi napovedni interval, kjer je:

$$c_\alpha^{nap} := F^{-1}_{Student(n-1)}\left(1-\frac{\alpha}{2}\right)\hat{\sigma}_+\sqrt{1+\frac{1}{n}}$$

4) Uporaba verjetja:

Deluje le za $Y = y = g(\theta)$, torej za količine, ki jih ne modeliramo kot slučajne.

Deluje dobro za veliko neodvisnih in enako porazdeljenih vrednosti.

Za verjetje sicer ni nujno, da so opažanja NEP, če pa imamo NEP, pa verjetje še toliko bolj pomaga.

Spomnimo se izreka 6.4:

Za cenilko $\hat{Y}$ po metodi največjega verjetja velja $MSE_{\theta}(\hat{Y}|Y) \sim \frac{\nabla g(\theta)^T (FI_1(\theta))^{-1} \nabla g(\theta)}{n}$.

Torej, več kot imamo informacij, manjša je napaka.

Izrek 8.1: Naj bo $\hat{\underline{\theta}}$ cenilka za $\underline{\theta}^* \in \Theta \overset{odp}{\subseteq} \mathbb{R}^n$ po metodi največjega verjetja na podlagi neodvisnih in enako porazdeljenih opažanj $X_1, \dots, X_n$ (kjer je nedvoumno definirana, sicer pa poljubna).
Naj bo $\underline{\theta}^* \in \Theta$ in $FI_1(\underline{\theta}^*)$ obrnljiva. Tedaj pod podobnimi dodatnimi pogoji kot za izrek 6.4 velja:

$$\sqrt{n}(\hat{\underline{\theta}} - \underline{\theta}^*) \xrightarrow[n \to \infty]{d,\, P_{\theta^*}} N(0, (FI_1(\theta^*))^{-1})$$

Nadalje, naj bo še $g : \Theta \to \mathbb{R}$ dovolj gladka funkcija ter $Y^* = g(\underline{\theta}^*)$ in $\hat{Y} = g(\hat{\underline{\theta}})$. Če definiramo:

$$RMSE^* := \left(\frac{(\nabla g(\theta^*))^T (FI_1(\theta^*))^{-1} (\nabla g(\theta^*))}{n}\right)^{1/2}$$

$$\widehat{RMSE} := \left(\frac{(\nabla g(\hat{\theta}))^T (FI_1(\hat{\theta}))^{-1} (\nabla g(\hat{\theta}))}{n}\right)^{1/2}$$

Velja še:

$$\frac{\hat{Y} - Y^*}{RMSE^*} \xrightarrow[n \to \infty]{d,\, P_{\theta^*}} N(0,1)$$

$$\frac{\hat{Y} - Y^*}{\widehat{RMSE}} \xrightarrow[n \to \infty]{d,\, P_{\theta^*}} N(0,1)$$

Če je torej $\hat{Y} = g(\hat{\theta}) = h(X_1, \dots, X_n)$ in $\widehat{RMSE} = s(X_1, \dots, X_n)$, potem je $\rho(x_1, \dots, x_n; y) = \frac{h(x_1, \dots, x_n) - y}{s(x_1, \dots, x_n)}$ asimptotično pristna funkcija.

Dokaz (ideje):

$$\nabla \ell(\underline{\theta} | x_1, \dots, x_n) \overset{NEP}{=} \sum_{i=1}^{n} \nabla \ell_1(\underline{\theta} | x_i)$$

Uporabimo večrazsežni CLI za vsoto $\sum_{i=1}^{n} \nabla l_1(\underline{\theta} | x_i)$.

Cenilka $\hat{y}$ je približno linearna funkcija vsote $\sum_{i=1}^{n} \nabla l_1(\underline{\theta}^* | x_i)$.

Uporabimo izrek o konvergenci slučajnih spremenljivk (trditev 2.21) in izrek Sluckega za deljenje (trditev 2.28).

$\Rightarrow$ Dobimo asimptotično eksakten interval zaupanja:

$$\hat{y} - \Phi^{-1}(1-\tfrac{\alpha}{2})\,\hat{RMSE} < y < \hat{y} + \Phi^{-1}(1-\tfrac{\alpha}{2})\,\hat{RMSE} \quad \in O\left(\tfrac{1}{\sqrt{n}}\right)$$

PREIZKUŠANJE DOMNEV, TESTIRANJE HIPOTEZ

Včasih imamo o θ kakšno "idejo". Imenujemo jo ničelna domneva oziroma hipoteza in jo označujemo s H_0.

Primer: "Kovanec je pošten"

$\theta = P(\text{pade grb}) \in [0,1]$

$H_0: \theta = \frac{1}{2}$

H_0 je v splošnem $\theta \in \Theta_0 \subseteq \Theta$.

Negaciji ničelne domneve H_0 pravimo alternativna domneva ali hipoteza in jo označimo s H_1.

Če je opažanje preveč neskladno s H_0, to domnevo zavrnemo in sprejmemo H_1.
Če pa opažanje ni neskladno, pa ne rečemo nič.

Torej do zmote po definiciji pride samo, če H_0 zavrnemo, čeprav velja.

Zato želimo H_0 zavrniti v čim več primerih, ko ne drži (in hkrati v čim manj primerih, ko ne drži).

Moč preizkusa je funkcija:

$$\Theta \setminus \Theta_0 \longrightarrow [0,1]$$
$$\theta \longmapsto P_\theta(H_0 \text{ zavrnemo})$$

Za $\theta \in \Theta_0$ pa zahtevamo:

stopnja tveganja

$$P_\theta(H_0 \text{ zavrnemo}) \ \square \ \alpha$$

Kjer je $\square$ ena izmed relacij:

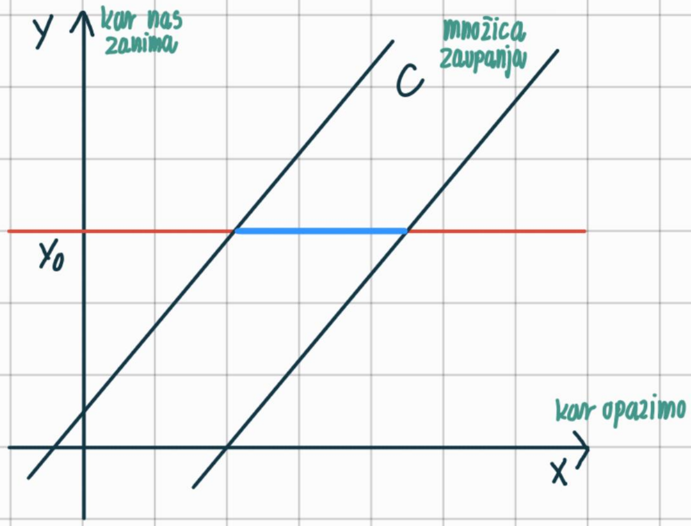
- $=$ (eksakten)
- $\approx$ (aproksimativen)
- $\xrightarrow[n\to\infty]{}$ (asimptotičen)
- $\leq$ (konservativen)

α imenujemo stopnja tveganja ali značilnosti preizkusa (significance level).

Zveza z množicami zaupanja:

Vsaka domneva $\theta \in \Theta_0$ je za primerno karakteristiko $Y = g(\theta)$ ekvivalentna domnevi $Y = Y_0$, na primer $g(\theta) = \mathbb{1}(\theta \in \Theta_0)$, $Y_0 = 1$.

Opazka 8.2: Nabor preizkusov domnev $H_0: Y = Y_0$ za vse možne Y_0 je ekvivalenten množici zaupanja za Y:



Domnevo $y = y_0$ zavrnemo za x, ki so izven C.

$\Rightarrow$ Če imamo nabor preizkusov za y, pogledamo navpične preseke s C.

$H_0: y = y_0$ zavrnemo $\Leftrightarrow (X, y_0) \notin C \Leftrightarrow y_0 \in C_X$

$\Leftrightarrow \rho(X, y_0) \geq c_\alpha$

mera razhajanja

kritična vrednost

Primer: Opazimo $S \sim Bin(n, \theta)$, kjer je n znan, θ pa nas zanima.

$\hat{\theta} = \frac{S}{n}$ (delež uspelih)

$\Rightarrow \rho(s, \theta) = \left|\frac{s}{n} - \theta\right|$

Hočemo, da je to pivotna funkcija, torej da ima stabilno porazdelitev.

$\frac{S}{n} - \theta = \hat{\theta} - \theta \overset{CLI}{\sim} N\left(0, \frac{\theta(1-\theta)}{n}\right)$

Ker θ tukaj nastopa, to ni pivotna funkcija. Niti blizu temu ni.

Lahko pa θ v varianci nadomestimo z oceno $\hat{\theta}$ in dobimo Waldov interval zaupanja:

$$\hat{\theta} - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}} < \theta < \hat{\theta} + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$$

(Abraham Wald, 1902-1950)

Izkaže se, da je to slab interval, saj nam diskretnost dela težave.

$H_0: \theta = \theta_0$ zavrnemo proti $H_1: \theta \neq \theta_0$, če velja:

$$|\hat{\theta} - \theta_0| \geq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$$

Naravnejša in tudi boljša različica preizkusa pa je, da $H_0: \theta = \theta_0$ zavrnemo, če velja:

$$|\hat{\theta} - \theta_0| \geq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\sqrt{\frac{\theta_0(1-\theta_0)}{n}}$$

To vodi v izboljšano množico zaupanja:

$$C = \left\{(s, \theta)\ ;\ \left|\frac{s}{n} - \theta\right| < \underbrace{\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}_{=:z}\sqrt{\frac{\theta(1-\theta)}{n}}\right\}$$

Pri danem s oziroma $\hat{\theta} = \frac{s}{n}$ dobimo:

$$\frac{2n\hat{\theta} + z^2 - \sqrt{4nz^2\hat{\theta}(1-\hat{\theta}) + z^4}}{2(n+z^2)} < \theta < \frac{2n\hat{\theta} + z^2 + \sqrt{4nz^2\hat{\theta}(1-\hat{\theta}) + z^4}}{2(n+z^2)}$$

To je Wilsonov interval zaupanja.

(Edwin Bidwell Wilson, 1879-1964)

Ta interval ima boljšo pokritost.

V splošnem lahko konstrukcijo množice zaupanja za karakteristiko y izboljšamo tako, da vzamemo:

$$C := \{(x, y)\ ;\ \rho(x, y) < c_\alpha(y)\}$$

Tukaj ni nujno, da je ρ pivotna. Dovolj je že, da je porazdelitev $\rho(X, y)$ pri vsakem θ natanko določena z $y = g(\theta)$.

Včasih nas skrbijo odstopanja le v eno smer:

$H_0: y = y_0$ preizkušamo proti $H_1^+: y > y_0$ ali $H_1^-: y < y_0$.

Primer: Loterija trdi, da vsaka druga srečka zadane.

Opazimo $S \sim Bin(n, \theta)$. Postavimo $\theta_0 = \frac{1}{2}$ in dobimo $H_0: \theta = \theta_0$ in $H_1^-: \theta < \theta_0$.

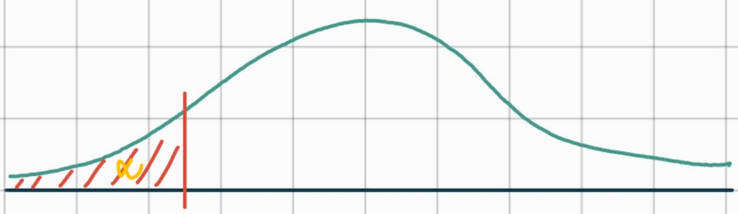
Če je n dovolj velik, pri H_0 velja $S \sim N(n\theta_0, \sqrt{n\theta_0(1-\theta_0)})$.

$$\hat{\theta} = \frac{s}{n} \sim N\left(\theta_0, \sqrt{\frac{\theta_0(1-\theta_0)}{n}}\right)$$

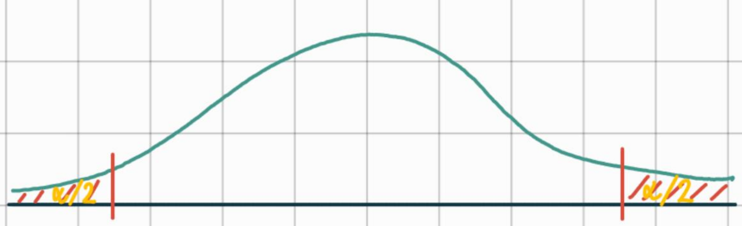
$$\Rightarrow \frac{\hat{\theta} - \theta_0}{\sqrt{\frac{\theta_0(1-\theta_0)}{n}}} \sim N(0,1)$$

$\Rightarrow H_0$ zavrnemo, če je $\frac{\hat{\theta} - \theta_0}{\sqrt{\frac{\theta_0(1-\theta_0)}{n}}} \leq -\Phi^{-1}(1-\alpha) = \Phi^{-1}(\alpha)$.

Enostransko območje zavrnitve:



Dvostransko območje zavrnitve:



PREIZKUŠANJE Z RAZMERJEM VERJETIJ

$$LR := \frac{\sup_{\theta \in \Theta} L(\theta|X)}{\sup_{\theta \in \Theta_0} L(\theta|X)} \qquad \text{(likelihood ratio)}$$

$H_0: \theta \in \Theta_0$ zavrnemo, če je $LR \geq c_\alpha$.

Prag c_α določimo skladno s stopnjo tveganja α.

(Samuel Stanley Wilks, 1906-1964)

Izrek 8.3 (Wilksov izrek):

Naj bodo $X_1, X_2, \ldots$ pri vsaki verjetnostni meri P_θ, $\theta \in \Theta$, neodvisne in enako porazdeljene slučajne spremenljivke, Θ naj bo gladka mnogoterost dimenzije p, $\Theta_0 \subseteq \Theta$ pa podmnogoterost dimenzije q. Obe naj bosta brez roba. Pod določenimi dodatnimi pogoji gre tedaj $2 \ln LR \xrightarrow[n \to \infty]{d} \chi^2(p-q)$ pri vseh P_θ, $\theta \in \Theta_0$.

Dokaz opustimo.

Wilksov preizkus:

$H_0: \theta \in \Theta_0$ zavrnemo, če je $2 \ln LR \geq F^{-1}_{\chi^2(p-q)}(1-\alpha)$.

Primer: $P_{\underline{\theta}} = \begin{pmatrix} a_0 & a_1 & \cdots & a_p \\ \theta_0 & \theta_1 & \cdots & \theta_p \end{pmatrix}$

$\underline{\theta} = (\theta_0, \theta_1, \ldots, \theta_p)$

$\Theta = \{ \underline{\theta} \in \mathbb{R}^{p+1} ; \theta_0, \theta_1, \ldots, \theta_p > 0, \sum_{j=1}^{p} \theta_j = 1 \}$

Za fiksne $\pi_0, \pi_1, \ldots, \pi_p$ preizkusimo:

$H_0: \theta_0 = \pi_0, \theta_1 = \pi_1, \ldots, \theta_p = \pi_p$

Opazimo:

$$\underbrace{X_1, X_2, \dots, X_p}_{\text{neodvisne}} \sim \begin{pmatrix} a_0 & a_1 & \cdots & a_p \\ \theta_0 & \theta_1 & \cdots & \theta_p \end{pmatrix}$$

$$\ln LR = \sup_{\theta \in \Theta} \sum_{i=1}^{n} l_1(\theta | X_i) - \sup_{\theta \in \Theta_0} \sum_{i=1}^{n} l_1(\theta | X_i)$$

$$l_1(\underline{\theta} | x) = \begin{cases} \ln\theta_0 & ; \ x = a_0 \\ \ln\theta_1 & ; \ x = a_1 \\ \vdots & \\ \ln\theta_p & ; \ x = a_p \end{cases}$$

$$= \sum_{j=0}^{p} \ln \theta_j \cdot \mathbb{1}(x = a_j)$$

$$l(\theta_0, \theta_1, \dots, \theta_p | X_1, X_2, \dots, X_n) = \sum_{i=1}^{n} \sum_{j=0}^{p} \ln \theta_j \cdot \mathbb{1}(X_i = a_j)$$

$$= \sum_{j=0}^{p} \ln \theta_j \underbrace{\sum_{i=1}^{n} \mathbb{1}(X_i = a_j)}_{N_j \text{ - frekvenca vrednosti } a_j}$$

Lagrange: $F = \sum_{j=0}^{p} N_j \ln\theta_j - \lambda \sum_{j=0}^{p} \theta_j$

$$\frac{\partial F}{\partial \theta_j} = \frac{N_j}{\theta_j} - \lambda = 0$$

$$\hat{\theta}_j = \frac{N_j}{\lambda}$$

$$\sum_{j=0}^{p} \hat{\theta}_j = 1 \Rightarrow \lambda = \sum_{j=0}^{p} N_j = n$$

$$\sup_{\theta \in \Theta} l(\theta_0, \theta_1, \dots, \theta_p | X_1, X_2, \dots, X_n) = \sum_{j=0}^{p} N_j \ln \hat{\theta}_j$$

$$\sup_{\theta \in \Theta} l(\theta_0, \theta_1, \dots, \theta_p | X_1, X_2, \dots, X_n) = \sum_{j=0}^{p} N_j \ln \pi_j$$

$$2 \ln LR = 2 \sum_{j=0}^{p} N_j \ln \frac{\hat{\theta}_j}{\pi_j} = 2 \sum_{j=0}^{p} N_j \ln \frac{N_j}{n\pi_j}$$

$\tilde{N}_j = n\pi_j$ so pričakovane frekvence.

Minimalni pogoji za legitimnost:

$$\forall j: \tilde{N}_j > 5$$

Tedaj bo porazdelitev $2\ln LR$ vsaj približno blizu $\chi^2(p)$.

Opazka 8.4: Izkaže se:

$$2\ln LR \approx \sum_{j=0}^{p} \frac{(N_j-\tilde{N}_j)^2}{\tilde{N}_j}$$

To je Pearsonov preizkus.

(Karl Pearson, 1857-1936)