Neyman-Pearson Theorem, Uniformly Most Powerful Test

A. Stepanov
D. Simpson
S. Culpepper

Consider a simple vs. simple hypothesis,

$$H_0\colon \theta = \theta_0 \quad vs. \quad H_1\colon \theta = \theta_1$$

Likelihood Ratio:

$$\Lambda(\mathbf{x}) = \frac{L(\theta_0;\mathbf{x})}{L(\theta_1;\mathbf{x})}$$

$$\alpha = \text{Type I Error} = P(Reject\ H_0 | H_0\ true)$$
$$\beta = \text{Power} = P(Reject\ H_0 | H_1\ true)$$

Neyman-Pearson Theorem:

$$\mathrm{C} = \{\mathbf{x} = (x_1, \dots, x_n)\colon \Lambda(\mathbf{x}) \le k\}$$

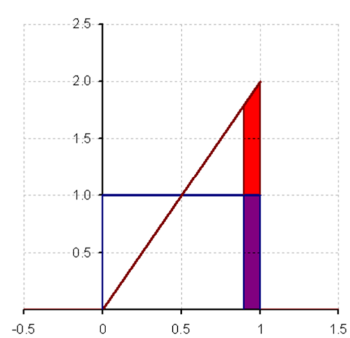
Reject H_0 if $\Lambda(\mathbf{x}) \le k$ is the best (most powerful) rejection region.

Example 1. Consider a family of probability distributions with pdf of the form $f(x;\theta) = \theta x^{\theta-1}, 0 < x < 1$, zero elsewhere, $\theta \ge 1$.

a) Test $H_0\colon \theta = 1$ vs. $H_1\colon \theta = 2$ with a sample X of size $n = 1$.

$$H_0\colon X \text{ has pdf } f(x;1) = \mathbf{1}_{0<x<1} \qquad vs. \qquad H_1\colon X \text{ has pdf } f(x;2) = 2x\,\mathbf{1}_{0<x<1}$$

$$\Lambda(x) = \frac{L(1;\mathbf{x})}{L(2;\mathbf{x})} = \frac{1}{2x} \le k \Leftrightarrow x \ge c = \frac{1}{2k}$$



The Type I error rate is,

$$\alpha = P_0(\Lambda(x) \le k) = P_0(X \ge c)$$
$$= \int_c^1 dx = 1 - c$$
$$\Rightarrow c = 1 - \alpha$$

The best (most powerful) rejection region is,

$$\mathrm{C} = \{x\colon \Lambda(x) \le k\} = \{x\colon x \ge c = 1 - \alpha\}.$$

Power is,

$$\beta = P_1(\Lambda(\mathbf{x}) \le k) = P_1(X \ge c) = \int_c^1 2x dx = 1 - c^2 = 1 - (1-\alpha)^2$$
$$= \alpha(2-\alpha)$$

b) Use a sample X_1, X_2 of size $n = 2$ to test $H_0: \theta = 1\ vs.\ H_1: \theta = \theta_1$.

$$\Lambda(x_1, x_2) = \frac{L(1; x_1, x_2)}{L(\theta_1; x_1, x_2)} = \frac{1}{\theta_1^2 (x_1 x_2)^{\theta_1 - 1}} \le k$$

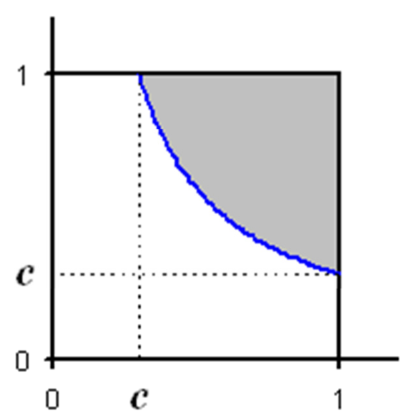
$$\Leftrightarrow x_1 x_2 \ge c = \left(\frac{1}{\theta_1^2 k}\right)^{\frac{1}{\theta_1 - 1}}$$

Best (most powerful) rejection region is given by

$$\text{C} = \{x_1, x_2 : \Lambda(x_1, x_2) \le k\} = \{x_1, x_2 : x_1 x_2 \ge c\}.$$

The Type I error rate as a function of c is,

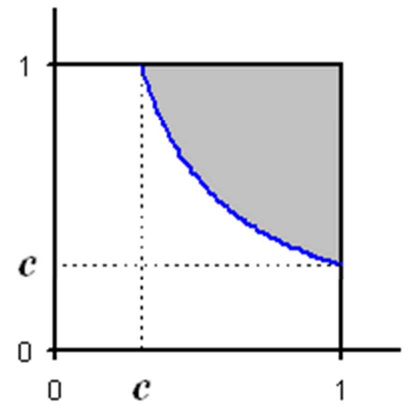
$$\alpha = P_0(\Lambda(x_1, x_2) \le k) = P_0(X_1 X_2 \ge c)$$
$$= \int_c^1 \int_{\frac{c}{x_1}}^1 dx_2 dx_1$$
$$= \int_c^1 \left(1 - \frac{c}{x_1}\right) dx_1$$
$$= (x_1 - c \ln x_1)|_c^1$$
$$= 1 - c + c \ln c$$



We would need to use a computer to solve for values of c given α.

Statistical power is,

$$\begin{aligned}\beta &= P_1(\Lambda(x_1, x_2) \le k) = P_1(X_1X_2 \ge c) \\ &= \theta_1^2 \int_c^1 \int_{\frac{c}{x_1}}^1 x_1^{\theta_1-1} x_2^{\theta_1-1}\, dx_2 dx_1 \\ &= \theta_1 \int_c^1 x_1^{\theta_1-1}\left(1 - \frac{c^{\theta_1}}{x_1^{\theta_1}}\right) dx_1 \\ &= \theta_1 \int_c^1 \left(x_1^{\theta_1-1} - \frac{c^{\theta_1}}{x_1}\right) dx_1 \\ &= 1 - c^{\theta_1} + \theta_1 c^{\theta_1} \ln c\end{aligned}$$



Example 2. Reconsider Example 1.

a) Use a sample $X_1, \dots, X_n$ to test $H_0\colon \theta = 1\ vs.\ H_1\colon \theta = \theta_1$.

$$\Lambda(\mathbf{x}) = \frac{L(1; \mathbf{x})}{L(\theta_1; \mathbf{x})} = \frac{1}{\theta_1^n \left(\prod_i^n x_i\right)^{\theta_1 - 1}} \le k$$

$$\prod_{i=1}^n x_i \ge c = \left(\frac{1}{\theta_1^n k}\right)^{\frac{1}{\theta_1-1}} \Leftrightarrow -\sum_{i=1}^n \ln x_i \le \tilde{c} = \left(\frac{n}{\theta_1 - 1}\right) \ln(\theta_1 k)$$

Best (most powerful) rejection region is given by

$$\mathrm{C} = \{\mathbf{x}\colon \Lambda(\mathbf{x}) \le k\} = \left\{\mathbf{x}\colon \prod_{i=1}^n x_i \ge c\right\} = \left\{\mathbf{x}\colon -\sum_{i=1}^n \ln x_i \le \tilde{c}\right\}.$$

Intuition:
1) In general, sums are easier to work with than the products. Logarithm turns a product into a sum.
2) If $0 < x < 1$, then $\ln x < 0$. Would rather work with positive random variables. So consider $-\ln x$.

Let $Y_i = -\ln X_i$ and recall $F_X(x) = x^\theta, 0 < x < 1$.

$$F_Y(y) = P(Y \le y) = P(-\ln X \le y) = P(X \ge e^{-y}) = 1 - e^{-\theta y}$$

$$Y_1, \dots, Y_n \; iid \; exponential(\lambda = \theta)$$

$$-\sum_{i=1}^{n} \ln X_i = \sum_{i=1}^{n} Y_i \sim Gamma\left(\alpha = n, "\theta" = \frac{1}{\theta}\right)$$

$$W = 2\theta \sum_{i=1}^{n} Y_i \sim \chi^2(2n)$$

So reject H_0 if $W \le 2\theta\tilde{c}$.

b) Find the Type I error rate.

The Type I error rate is,

$$\alpha = P_0(\Lambda(\mathbf{x}) \le k) = P_0(W \le 2\tilde{c})$$
$$\Rightarrow \tilde{c} = \frac{\chi^2_{1-\alpha}(2n)}{2}$$

c) Find power.

The power function is,

$$\beta = P_1(\Lambda(\mathbf{x}) \le k) = P_1(W \le 2\theta_1\tilde{c}) = P_1\left(W \le \theta_1\chi^2_{1-\alpha}(2n)\right)$$

OR

The Gamma-Poisson relationship can be used.

If $T \sim Gamma(k, \theta = 1/\lambda)$ where k is an integer, then

$$P(T_k \ge t) = P(X_t \le k - 1)$$

where $X_t \sim Poisson(\lambda t = t/\theta)$.

Recall $W \sim Gamma\left(k = \frac{2n}{2} = n, \theta = 2\right)$.

$$P_1\left(W \leq \theta_1 \chi^2_{1-\alpha}(2n)\right) = P_1\left(T_n \leq t = \theta_1 \chi^2_{1-\alpha}(2n)\right) = P(X_t \geq n)$$

where $X_t \sim Poisson\left(\frac{\theta_1 \chi^2_{1-\alpha}(2n)}{2}\right)$.

d) Suppose $n = 10$. Test $H_0: \theta = 1\ vs.\ H_1: \theta = 2$ with a significance level of $\alpha = .05$. Find power.

$$\tilde{c} = \frac{\chi^2_{0.95}(20)}{2} = \frac{10.85}{2} = 5.425$$

From R, pchisq(2*10.85,20) yields,

$$\beta = P_1(W \leq 2 \cdot 10.85) = 0.643$$

OR

$$\beta = P_1(W \leq 2 \cdot 10.85) = P_1(T_{10} \leq 2 \cdot 10.85) = P(X_{2 \cdot 10.85} \geq 10)$$

where $X_t \sim Poisson\left(2 \cdot \frac{10.85}{2} = 10.85\right)$.

Using the 11.0 column for the Poisson cumulative table implies,

$$P[Poisson(10.85) \geq 10] \approx P[Poisson(11) \geq 10] = 1 - 0.341 = 0.659$$

Example 3. Let $\lambda > 0$ and let $X_1, \ldots, X_n$ be a random sample from the distribution with the probability density function,

$$f(x;\lambda) = 2\lambda^2 x^3 e^{-\lambda x^2}, \qquad x > 0$$

We wish to test: $H_0: \lambda = 5 \;\; vs. \;\; H_1: \lambda = 3$.

a) Find the form of the most powerful rejection region.

$$L(\lambda;\mathbf{x}) = 2^n \lambda^{2n} \exp\left(-\lambda \sum_{i=1}^n x_i^2\right)\left(\prod_{i=1}^n x_i\right)^3$$

$$\Lambda(\mathbf{x}) = \frac{L(5;\mathbf{x})}{L(3;\mathbf{x})} = \frac{2^n 25^n \exp\left(-5\sum_{i=1}^n x_i^2\right)\left(\prod_{i=1}^n x_i\right)^3}{2^n 9^n \exp\left(-3\sum_{i=1}^n x_i^2\right)\left(\prod_{i=1}^n x_i\right)^3}$$
$$= \left(\frac{25}{9}\right)^n \exp\left(-2\sum_{i=1}^n x_i^2\right) \le k$$

$$\Lambda(\mathbf{x}) \le k \Leftrightarrow \sum_{i=1}^n x_i^2 \ge c = -\frac{1}{2}\ln\left[\left(\frac{9}{25}\right)^n k\right]$$
$$\mathrm{C} = \{\mathbf{x}: \Lambda(\mathbf{x}) \le k\} = \left\{\mathbf{x}: \sum_{i=1}^n x_i^2 \ge c\right\}.$$

Recall,

$$\sum_{i=1}^n X_i^2 \sim Gamma\left(\alpha = 2n, \theta = \frac{1}{\lambda}\right).$$

So,

$$W = 2\lambda \sum_{i=1}^n x_i^2 \sim \chi^2(4n)$$

Reject H_0 if $W \ge \chi^2_\alpha(4n) = 2\lambda c \Rightarrow c = \chi^2_\alpha(4n)/2\lambda$.

b) Suppose $n = 4$. Find the most powerful rejection region with a 5% level of significance.

$$\chi^2_{0.05}(16) = 26.30 \Rightarrow c = \frac{26.30}{10} = 2.63$$

c) Consider the rejection region "Reject H_0 if $\sum_{i=1}^{4} x_i^2 \geq 2.5$". Find the significance level of this test.

$$\alpha = P_0\left(\sum_{i=1}^{4} X_i^2 \geq 2.5\right) = P(T_8 \geq 2.5) = P(X_{2.5} \leq 7)$$
$$= P(Poisson(5 \cdot 2.5) \leq 7) = 0.070$$

OR

$$\sum_{i=1}^{4} x_i^2 \geq 2.5 \Leftrightarrow W \geq 5\lambda$$

$$P_0(W \geq 25) = P(T_8 \geq 25) = P(X_{25} \leq 7) = P\left(Poisson\left(\frac{25}{2}\right) \leq 7\right) = 0.070$$

d) Consider the rejection region "Reject H_0 if $\sum_{i=1}^{4} x_i^2 \geq 2.5$". Find the power of this test.

$$\beta = P_1\left(\sum_{i=1}^{4} X_i^2 \geq 2.5\right) = P(X_{2.5} \leq 7) = P(Poisson(3 \cdot 2.5) \leq 7) = 0.525$$

OR

$$\beta = P_1(W \geq 15) = P(X_{15} \leq 7) = P\left(Poisson\left(\frac{15}{2}\right) \leq 7\right) = 0.525$$

e) Recall the best most powerful test is "Reject H_0 if $\sum_{i=1}^{n} x_i^2 \geq c$". Consider an alternative rejection region "Reject H_0 if $\sum_{i=1}^{n} x_i^2 \leq 0.8$". Find the significance and power of the test.

$$\alpha = P_0\left(\sum_{i=1}^{4} X_i^2 \leq 0.8\right) = P(T_8 \leq 0.8) = P(X_{0.8} \geq 8)$$
$$= 1 - P(Poisson(5 \cdot 0.8) \leq 7) = 0.051$$

$$\beta = P_1\left(\sum_{i=1}^{4} X_i^2 \leq 0.8\right) = P(X_{0.8} \geq 7) = 1 - P(Poisson(3 \cdot 0.8) \leq 7)$$
$$= 1 - 0.997 = 0.003$$

Clearly, the alternative rejection region is not the most powerful!

f) Suppose $\sum_{i=1}^{4} x_i^2 = 3.2$. Find the p-value of this test.

$$p - value = P_0\left(\sum_{i=1}^{4} x_i^2 \geq 3.2\right) = P(T_8 \geq 3.2) = P(X_{3.2} \leq 7)$$
$$= P(Poisson(5 \cdot 3.2) \leq 7) = 0.010$$

Neyman-Pearson Theorem, Uniformly Most Powerful Test

A. Stepanov
D. Simpson
S. Culpepper

Definition. The critical region C is a **uniformly most powerful (UMP) critical region** of size α for testing the simple hypothesis H_0 versus an alternative composite hypothesis H_1 if the set C is a best critical region of size α for testing H_0 against each simple hypothesis in H_1.

Example 4. Suppose $X \sim Exp(\theta)$ and consider testing the simple versus composite hypothesis,

$$H_0: \theta = 1 \; vs. H_1: \theta > 1$$

Is there a UMP?

We showed the LRT evaluated at $\theta = 1$ versus $\theta = A > 1$ is

$$\Lambda(x) = \frac{f_0(x)}{f_1(x)} = \frac{e^{-x}}{\frac{1}{A} e^{-\frac{x}{A}}} = Ae^{-\frac{A-1}{A}x} \le k$$

The best critical region for a given A is,

$$\Lambda(x) \le k \quad \Leftrightarrow \quad X \ge c = \frac{A}{A-1} \ln\left(\frac{A}{k}\right)$$

We showed

$$c = -\ln \alpha$$

which implies,

$$\mathrm{C} = \{x: \Lambda \le k\} = \{x: x \ge -\ln(\alpha)\}$$

Notice that C is a best critical region of size α for testing H_0 against each simple hypothesis in $H_1: \theta > 1$.

C is UMP.

Example 5. Suppose $X \sim Exp(\theta)$ and consider testing the simple versus composite hypothesis,

$$H_0: \theta = 1 \; vs. H_1: \theta \neq 1$$

Is there a UMP?

$$\Lambda(x) = \frac{f_0(x)}{f_1(x)} = \frac{e^{-x}}{\frac{1}{A}e^{-\frac{x}{A}}} = Ae^{-\frac{A-1}{A}x} \leq k$$

The best critical region for a given A is,

$$\Lambda(x) \leq k \quad \Leftrightarrow \begin{cases} X \leq c = \dfrac{A}{1-A}\ln\left(\dfrac{k}{A}\right) & A < 1 \\ X \geq c = \dfrac{A}{A-1}\ln\left(\dfrac{A}{k}\right) & A > 1 \end{cases}$$

The critical value for a given α is,

$$\alpha = \begin{cases} P_0(X \leq c) & A < 1 \\ P_0(X \geq c) & A > 1 \end{cases} = \begin{cases} 1 - e^{-c} & A < 1 \\ e^{-c} & A > 1 \end{cases}$$

The LRT critical value is,

$$c = \begin{cases} -\ln(1-\alpha) & A < 1 \\ -\ln \alpha & A > 1 \end{cases}$$

$$\mathrm{C} = \{x: \Lambda \leq k\} = \left\{x: \begin{matrix} x \leq -\ln(1-\alpha) & A < 1 \\ x \geq -\ln(\alpha) & A > 1 \end{matrix}\right\}$$

The best and most powerful confidence region, C, is not uniform for all values of θ under H_1.

Example 6. Let $X_1, \dots, X_n$ be an iid sample from an exponential distribution with mean θ,

$$f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \qquad x > 0$$

We wish to test $H_0: \theta = \frac{1}{2}$ $vs.$ $H_1: \theta > \frac{1}{2}$.

a) If $n = 7$, find a UMP rejection region with significance level $\alpha = 0.05$.

$$\Lambda(\mathbf{x}) = \frac{L\left(\frac{1}{2};\mathbf{x}\right)}{L(\theta;\mathbf{x})} = \frac{\frac{1}{2^7}\exp(-2\sum_{i=1}^7 x_i)}{\frac{1}{\theta^7}\exp\left(-\frac{1}{\theta}\sum_{i=1}^7 x_i\right)} = \left(\frac{\theta}{2}\right)^7 \exp\left(-\left(2-\frac{1}{\theta}\right)\sum_{i=1}^7 x_i\right) \le k$$

$$\Lambda(\mathbf{x}) \le k \Leftrightarrow \sum_{i=1}^7 x_i \ge c = -\frac{\theta}{2\theta - 1}\ln\left(k\left(\frac{2}{\theta}\right)^7\right)$$

$$\sum_{i=1}^7 X_i \sim Gamma\left(\alpha = 7, \theta = \frac{1}{\lambda}\right)$$

$$Y = \frac{2\sum_{i=1}^7 X_i}{\theta} \sim \chi^2(14)$$

$$\alpha = P_0\left(Y \ge \frac{2c}{\frac{1}{2}}\right) \quad \Rightarrow \quad c = \frac{\chi^2_{0.05}(14)}{4} = \frac{23.68}{4} = 5.92$$

Reject H_0 if $\sum_{i=1}^7 x_i \ge 5.92$.

b) Find the power of the rejection rule from a) at $H_1: \theta = \frac{3}{4}$.

$$\beta = P_1\left(\sum_{i=1}^7 X_i \ge 5.92\right) = P_1(T_7 \ge 5.92) = P(X_{5.92} \le 6) > P(X_6 \le 6)$$
$$= P\left(Poisson\left(\frac{4}{3}6 = 8\right) \le 6\right) = 0.313$$

c) Find the significance level if the rejection rule is "Reject H_0 if $\sum_{i=1}^{7} x_i \geq 6$".

$$\alpha = P_0\left(\sum_{i=1}^{7} X_i \geq 6\right) = P_0(T_7 \geq 6) = P(X_6 \leq 6)$$
$$= P\left(Poisson\left(\frac{6}{1/2} = 12\right) \leq 6\right) = 0.046$$

d) Find the power if the rejection rule is "Reject H_0 if $\sum_{i=1}^{7} x_i \geq 6$" at $\theta = \frac{3}{4}, 1, \frac{3}{2}, 2$.

$$\beta = P_0\left(\sum_{i=1}^{7} X_i \geq 6\right) = P_0(T_7 \geq 6) = P(X_6 \leq 6) = P\left(Poisson\left(\frac{6}{\theta}\right) \leq 6\right)$$

$$\beta = \begin{cases} P\left(Poisson\left(\frac{4}{3}6 = 8\right) \leq 6\right) = 0.313 & \theta = \frac{3}{4} \\ P(Poisson(6) \leq 6) = 0.606 & \theta = 1 \\ P\left(Poisson\left(\frac{2}{3}6 = 4\right) \leq 6\right) = 0.889 & \theta = \frac{3}{2} \\ P(Poisson(3) \leq 6) = 0.996 & \theta = 2 \end{cases}$$