WIKIPEDIA

Uniformly most powerful test

In statistical hypothesis testing, a **uniformly most powerful** (**UMP**) **test** is a hypothesis test which has the **greatest power** $1-\beta$ among all possible tests of a given size α. For example, according to the Neyman–Pearson lemma, the likelihood-ratio test is UMP for testing simple (point) hypotheses.

Contents

Setting

Let X denote a random vector (corresponding to the measurements), taken from a parametrized family of probability density functions or probability mass functions $f_\theta(x)$, which depends on the unknown deterministic parameter $\theta \in \Theta$. The parameter space Θ is partitioned into two disjoint sets Θ_0 and Θ_1. Let H_0 denote the hypothesis that $\theta \in \Theta_0$, and let H_1 denote the hypothesis that $\theta \in \Theta_1$. The binary test of hypotheses is performed using a test function $\varphi(x)$ with a reject region R (a subset of measurement space).

$$\varphi(x) = \begin{cases} 1 & \text{if } x \in R \\ 0 & \text{if } x \in R^c \end{cases}$$

meaning that H_1 is in force if the measurement $X \in R$ and that H_0 is in force if the measurement $X \in R^c$. Note that $R \cup R^c$ is a disjoint covering of the measurement space.

Formal definition

A test function $\varphi(x)$ is UMP of size α if for any other test function $\varphi'(x)$ satisfying

$$\sup_{\theta\in\Theta_0} \operatorname{E}[\varphi'(X)|\theta] = \alpha' \leq \alpha = \sup_{\theta\in\Theta_0} \operatorname{E}[\varphi(X)|\theta]$$

we have

$$\forall \theta \in \Theta_1, \quad \operatorname{E}[\varphi'(X)|\theta] = 1 - \beta'(\theta) \leq 1 - \beta(\theta) = \operatorname{E}[\varphi(X)|\theta].$$

The Karlin–Rubin theorem

The Karlin–Rubin theorem can be regarded as an extension of the Neyman–Pearson lemma for composite hypotheses.[1] Consider a scalar measurement having a probability density function parameterized by a scalar parameter θ, and define the likelihood ratio $l(x) = f_{\theta_1}(x)/f_{\theta_0}(x)$. If $l(x)$ is monotone non-decreasing, in x, for any pair $\theta_1 \geq \theta_0$ (meaning that the greater x is, the more likely H_1 is), then the threshold test:

$$\varphi(x) = \begin{cases} 1 & \text{if } x > x_0 \\ 0 & \text{if } x < x_0 \end{cases}$$

where x_0 is chosen such that $\operatorname{E}_{\theta_0} \varphi(X) = \alpha$

is the UMP test of size α for testing $H_0 : \theta \leq \theta_0 \text{ vs. } H_1 : \theta > \theta_0$.

Note that exactly the same test is also UMP for testing $H_0 : \theta = \theta_0 \text{ vs. } H_1 : \theta > \theta_0$.

Important case: exponential family

Although the Karlin-Rubin theorem may seem weak because of its restriction to scalar parameter and scalar measurement, it turns out that there exist a host of problems for which the theorem holds. In particular, the one-dimensional exponential family of probability density functions or probability mass functions with

$$f_\theta(x) = g(\theta)h(x)\exp(\eta(\theta)T(x))$$

has a monotone non-decreasing likelihood ratio in the sufficient statistic $T(x)$, provided that $\eta(\theta)$ is non-decreasing.

Example

Let $X = (X_0, \ldots, X_{M-1})$ denote i.i.d. normally distributed N-dimensional random vectors with mean θm and covariance matrix R. We then have

$$\begin{aligned} f_\theta(X) &= (2\pi)^{-MN/2}|R|^{-M/2}\exp\left\{-\frac{1}{2}\sum_{n=0}^{M-1}(X_n - \theta m)^T R^{-1}(X_n - \theta m)\right\} \\ &= (2\pi)^{-MN/2}|R|^{-M/2}\exp\left\{-\frac{1}{2}\sum_{n=0}^{M-1}\left(\theta^2 m^T R^{-1} m\right)\right\} \\ &\quad \exp\left\{-\frac{1}{2}\sum_{n=0}^{M-1}X_n^T R^{-1} X_n\right\}\exp\left\{\theta m^T R^{-1}\sum_{n=0}^{M-1}X_n\right\} \end{aligned}$$

which is exactly in the form of the exponential family shown in the previous section, with the sufficient statistic being

$$T(X) = m^T R^{-1}\sum_{n=0}^{M-1}X_n.$$

Thus, we conclude that the test

$$\varphi(T) = \begin{cases} 1 & T > t_0 \\ 0 & T < t_0 \end{cases} \qquad \mathrm{E}_{\theta_0}\, \varphi(T) = \alpha$$

is the UMP test of size α for testing $H_0 : \theta \leqslant \theta_0$ vs. $H_1 : \theta > \theta_0$

Further discussion

Finally, we note that in general, UMP tests do not exist for vector parameters or for two-sided tests (a test in which one hypothesis lies on both sides of the alternative). The reason is that in these situations, the most powerful test of a given size for one possible value of the parameter (e.g. for θ_1 where $\theta_1 > \theta_0$) is different from the most powerful test of the same size for a different value of the parameter (e.g. for θ_2 where $\theta_2 < \theta_0$). As a result, no test is **uniformly** most powerful in these situations.

References

1. Casella, G.; Berger, R.L. (2008), *Statistical Inference*, Brooks/Cole. ISBN 0-495-39187-5 (Theorem 8.3.17)

Further reading

- Ferguson, T. S. (1967). "Sec. 5.2: *Uniformly most powerful tests*". *Mathematical Statistics: A decision theoretic approach*. New York: Academic Press.
- Mood, A. M.; Graybill, F. A.; Boes, D. C. (1974). "Sec. IX.3.2: *Uniformly most powerful tests*". *Introduction to the theory of statistics* (3rd ed.). New York: McGraw-Hill.
- L. L. Scharf, *Statistical Signal Processing*, Addison-Wesley, 1991, section 4.7.