Mathematical Definitions

Ofir Shukron

1 P-values

Given a rejection region Γ in the form of $[c, \infty)]$, for the null hypothesis H_0, the p-value of an observed statistics $T = t$ is defined as

$$p - value(t) = min_{\{\Gamma; t \in \Gamma\}} \{Pr(T \in \Gamma | H_0 \quad true)\}$$

See definition in [3].

2 Type I and Type II errors

A type I and type II errors might occur in statistical hypothesis testing when hypothesizing about the observed null hypothesis. Type I error is accepting the alternative (or rejecting the null) when the null is correct, while type II is accepting the null when the alternative is correct.

3 Quadratic Forms

From Mathai [2](Chapters 1 and 4). For a $p \times 1$ random vector X with mean value $E(X) = \mu$ and $Cov(X) = E[(X - E(X))(X - E(X))'] = \Sigma > 0$, we set

$Y = \Sigma^{-0.5} X \Rightarrow E(Y) = \Sigma^{-0.5}\mu$ and $Cov(Y) = \Sigma^{-0.5} Cov(X) \Sigma^{-0.5} = I$ $Z = (Y - \Sigma^{-0.5}\mu) \Rightarrow E(Z) = 0$ and $Cov(Z) = I$ We can express the quadratic form $Q(X) = X'AX$ with the centralize variable Z and a symmetric positive definite matrix A as

$$Q(X) = X'AX = Y'\Sigma^{0.5} A \Sigma^{0.5} Y = (Z + \Sigma^{-0.5}\mu)' \Sigma^{0.5} A \Sigma^{0.5} (Z + \Sigma^{-0.5}\mu) \quad (1)$$

As an example, if we set $\mu = 0$, $A = I$, $Cov(X) = \Sigma = \sigma I$ and get $Q(X) = X'X$, $Y = \frac{X}{\sqrt{\sigma}}$, $Z = (\frac{X}{\sqrt{\sigma}} - 0)$, then

$$Q(X) = X'X = (\frac{X}{\sqrt{\sigma}})'(\sigma I)(\frac{X}{\sqrt{\sigma}}) = XX' \tag{2}$$

which is the normal form.

If P is a $p \times p$ orthogonal matrix which diagonalize $\Sigma^{0.5} A \Sigma^{0.5}$, that is

$$P'\Sigma^{0.5} A \Sigma^{0.5} P = diag(\lambda_1, \lambda_2, ..., \lambda_p)$$

with $P'P = I$, and λ_i are the eigenvalues of $\Sigma^{0.5} A \Sigma^{0.5}$. Then if $U = P'Z$, we have

$$Z = PU, \qquad E(U) = 0 \qquad Cov(U) = I$$

Then we can express the quadratic form of X by

$$Q(X) = (U + b)' diag(\lambda_1, \lambda_2, ..., \lambda_p)(U + b) \tag{3}$$

with $b = (P'\Sigma^{-0.5}\mu)'$

4 Spline interpolation

This section is mainly taken from [1]. The Newton from of the interpolation polynomial is defined using the divided differences as:

$$p_n(x) = \sum_{i=1}^{n} (x - \tau_1) \cdots (x - \tau_n)[\tau_1, \tau_2, ..\tau_n]g$$

where τ_i are the nodes of the function g to be interpolated, in which the polynomial $p_n(x)$ of order n agrees with it i.e $p_n(\tau_i) = g(\tau_i) \quad \forall i = 1..n$

References

[1] Carl De Boor. A practical guide to splines. *Mathematics of Computation*, 1978.

[2] Arakaparampil M Mathai and Serge B Provost. *Quadratic forms in random variables: theory and applications*. M. Dekker New York, 1992.

[3] John D Storey. A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(3):479–498, 2002.