Introduction to Statistical Method

Simultaneous Estimation of the Mean and Variance

Chi Random Variable

Consider a problem:

- $z = (z_1, \cdots, z_n) \in \mathbb{R}^n$ randomly chosen.
- the value z_k is determined by random variable Z_k, following a standard normal distribution.
- think about distribution about $\chi_n := \sqrt{\sum_{i=1}^n Z_i^2}$
- χ_n is a chi random variable, follows chi distribution with n degree of freedom.

Cumulative Distribution Function

$$F_{\chi_n}(y) = P[\chi_n \le y] = P[\chi_n^2 \le y^2] = P[\sum_{i=1}^n Z_i^2 \le y^2] = \int_{\sum_{k=1}^n z_k^2 \le y^2} f_{Z_1 \dots Z_n}(z_1, \cdots, z_n) dz_1 \dots dz_n$$

Since they are $Z_1, \dots Z_n$ that n independent standard variables, then we see the joint density:

$$f_{Z_1 \dots Z_n}(z_1, \dots z_n) = \frac{1}{(2\pi)^{n/2}} e^{-\sum_{k=1}^n z_k^2/2}$$

Thus $F_{\chi_n}(y) = \int_{\sum_{k=1}^n z_k^2 \le y^2} (2\pi)^{-n/2} e^{-\sum_{k=1}^n z_k^2/2} dz_1 \dots dz_n$

Apply polar coordinate with $(r, \theta_1, \dots, \theta_0)$ with $r > 0$, $0 < \theta_{n-1} < 2\pi$ and $-\pi/2 < \theta_k < \pi/2$ for $k = 1, \dots n-2$:

$$\begin{aligned} x_1 &= r \sin\theta_1 \\ x_2 &= r \cos\theta_1 \sin\theta_2 \\ x_3 &= r \cos\theta_1 \cos\theta_2 \sin\theta_3 \\ &\vdots \\ x_{n-1} &= r \cos\theta_1 \cos\theta_2 \dots \cos\theta_{n-2} \sin\theta_{n-1} \\ x_n &= r \cos\theta_1 \cos\theta_2 \dots \cos\theta_{n-2} \cos\theta_{n-1} \end{aligned}$$

Then the integral become: $$\begin{aligned} F_{\chi_n}(y) = &\int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} \int_0^y (2\pi)^{-n/2} e^{-r^2/2}\, r^{n-1} \\ &\times D(\theta_1, \dots, \theta_{n-1})\, dr\, d\theta_1 \dots d\theta_{n-2}\, d\theta_{n-1} \end{aligned}$$

Since $D(\theta_1, \dots \theta_{n-1})$ is independent of r, $C_n = (2\pi)^{-n/2} \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} D(\theta_1, \dots \theta_{n-1}) d\theta_1 \cdots d\theta_{n-1}$

we have $F_{\chi_n}(y) = C_n \int_0^y e^{-r^2/2} r^{n-1} dr$.

Obviously, the $1 = \lim_{y \to \infty} C_n \int_0^\infty e^{-r^2/2} r^{n-1} dr = C_n \Gamma(\frac{n}{2}) 2^{n/2-1}$

Thus $C_n = (\Gamma(\frac{n}{2}) 2^{n/2-1})^{-1}$ and $f_{\chi_n}(y) = \frac{2}{\Gamma(\frac{n}{2}) 2^{n/2}} y^{n-1} e^{-y^2/2}$

Chi-Squared Distribution

we hence derive from the $F_{\chi^2_n} = (\Gamma(\frac{n}{2})2^{n/2-1})^{-1} \int_0^{\sqrt{y}} e^{-r^2/2} r^{n-1} dr$

$f_{\chi^2_n} = F'_{\chi^2_n}(y) = (\Gamma(\frac{n}{2})2^{n/2-1})^{-1} e^{-y/2} \sqrt{y}^{n-1} \cdot \frac{d}{dy}\sqrt{y} = \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} y^{n/2-1} e^{-y/2}$

Sum of Independent Chi-Squared Variables

Given $\chi^2_m = \sum_{i=1}^{m} X_i^2$ and $\chi^2_n = \sum_{j=1}^{n} Y_j^2$, then $\chi^2_{m+n} = \chi^2_m + \chi^2_n = \sum_{i=1}^{m} X_i^2 + \sum_{j=1}^{n} Y_j^2$

It follows a chi-squared distribution, but with $m + n$ degree of freedom.

It extends to multi-addition case, trivial.