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

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

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

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

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

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

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

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

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

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 $F_{\chi_n^2} = (\Gamma(\frac{n}{2})2^{n/2-1})^{-1}\int_0^{\sqrt{y}} e^{-r^2/2}r^{n-1}dr$

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

Sum of Independent Chi-Squared Variables

Given $\chi_m^2 = \sum_{i=1}^{m} X_i^2$ and $\chi_n^2 = \sum_{j=1}^{n} Y_j^2$, then $\chi_{m+n}^2 = \chi_m^2 + \chi_n^2 = \sum_{i=1}^{m} X_i^2 + \sum_{j=1}^{n} Y_j^2$

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

It extends to multi-addition case, trivial.

Joint Sampling of Mean and Variance

In the previous chapter, we were able to analyze the sample mean, and also its distribution, under the assumption of known variance.

If variance $\sigma^2 = E[(X-\mu)^2]$ is unknown, then we must first see $S^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X})^2$

So we are using the random sample $X_1, \dots X_n$ to get $\overline{X}$ and S^2 at same time.

So we are getting the joint distribution of $\overline{X}$ and S^2.

Theorem

Predicate

- $X_1, \dots X_n$ $n \geq 2$ be a random sample of size n.
- Normal distribution with μ and variance σ^2.

Content

- The sample mean $\overline{X}$ is independent of the sample variance S^2.
- $\overline{X}$ is normally distributed with mean μ and σ^2/n.
- $(n-1)S^2/\sigma^2$ is chi-squared distributed with $n-1$ degree of freedom.

Helmert Transformation

The Helmert transformation is a very special kind of orthogonal transformation from a set of $n \geq 2$ normal random variables $X_1, \dots X_n$ to a new set of random variables $Y_1, \dots Y_n$.

$$
\begin{aligned}
Y_1 &= \frac{1}{\sqrt{n}}(X_1 + \cdots + X_n) \\
Y_2 &= \frac{1}{\sqrt{2}}(X_1 - X_2) \\
Y_3 &= \frac{1}{\sqrt{6}}(X_1 + X_2 - 2X_3) \\
&\vdots \\
Y_n &= \frac{1}{\sqrt{n(n-1)}}(X_1 + X_2 + \cdots + X_{n-1} - (n-1)X_n)
\end{aligned}
$$

$$
\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & \cdots & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & -\frac{n-1}{\sqrt{n(n-1)}} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_n \end{pmatrix}
$$

The matrix A is orthonormal since $A^{-1} = A^T$. This implies $|\det A| = 1$.

$$
\sum_{i=1}^{n} y_i^2 = \langle y, y \rangle = \langle Ax, Ax \rangle = (Ax)^T(Ax) = x^T A^T Ax = \langle A^T Ax, x \rangle = \langle x, x \rangle = \sum_{i=1}^{n} x_i^2
$$

Thus the joint distribution:

$$
\begin{aligned}
f_{X_1\cdots X_n}(x_1,\dots,x_n) &= \prod_{i=1}^{n}(2\pi)^{-1/2}\sigma^{-1}e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}\\
&= (2\pi)^{-n/2}\sigma^{-n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i^2-2\mu x_i+\mu^2)}\\
&= (2\pi)^{-n/2}\sigma^{-n}e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}x_i^2-2\mu\sum_{i=1}^{n}x_i+n\mu^2\right)}
\end{aligned}
$$

Apply back into the y_n,

$$
\begin{aligned}
&f_{Y_1\cdots Y_n}(y_1,\dots,y_n)\\
&= f_{Y_1\cdots Y_n}(\mathbf{y}) = f_{X_1\cdots X_n}(\mathbf{x})_{\mathbf{x}=A^{-1}\mathbf{y}}\cdot\underbrace{|\det DA^{-1}(\mathbf{y})|}_{=1}\\
&= (2\pi)^{-n/2}\sigma^{-n}e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}y_i^2-2\mu\sqrt{n}y_1+n\mu^2\right)}\\
&= (2\pi)^{-n/2}\sigma^{-n}e^{-\frac{1}{2\sigma^2}\left(\sum_{i=2}^{n}y_i^2+(y_1-\sqrt{n}\mu)^2\right)}\\
&= (2\pi)^{-1/2}\sigma^{-1}e^{-\frac{1}{2\sigma^2}(y_1-\sqrt{n}\mu)^2}\prod_{i=2}^{n}(2\pi)^{-1/2}\sigma^{-1}e^{-\frac{1}{2\sigma^2}y_i^2}
\end{aligned}
$$

Then $f_{Y_1}(y_1)=(2\pi)^{-1/2}\sigma^{-1}e^{-\frac{1}{2\sigma^2}(y_1-\sqrt{n}\mu)^2}$ and $f_{Y_i}(y_i)=(2\pi)^{-1/2}\sigma^{-1}e^{-\frac{1}{2\sigma^2}y_i^2}$ for $2\le i\le n$.

$f_{Y_1\ldots Y_n}(y_1,\ldots y_n)=f_{Y_1}(y_1)\cdot f_{Y_2}(y_2)\ldots f_{Y_n}(y_n)$

So Y_1 is normally distributed with mean $\sqrt{n}\mu$ and variance σ^2, while $Y_2\ldots Y_n$ are having mean 0 and variance σ^2.

Proof for Previous Theorem

So $\overline{X}=n^{-1/2}Y_1$ and

$$
\begin{aligned}
(n-1)S^2 &= \sum_{k=1}^{n}(X_k-\overline{X})^2 = \sum_{k=1}^{n}X_k^2-2\sum_{k=1}^{n}X_k\overline{X}+n\overline{X}^2\\
&= \sum_{k=1}^{n}X_k^2-n\overline{X}^2 = \sum_{k=1}^{n}Y_k^2-Y_1^2 = \sum_{k=2}^{n}Y_k^2
\end{aligned}
$$

Since $\overline{X}=n^{-1/2}Y_1$ and $f_{Y_1}(y_1)=(2\pi)^{-1/2}\sigma^{-1}e^{-\frac{1}{2\sigma^2}(y_1-\sqrt{n}\mu)^2}$, then according to the rule in `ve401 note 3 page 6` that we get $f_{\overline{X}}(x)=(2\pi)^{-1/2}\sigma^{-1}e^{-\frac{1}{2\sigma^2}(\sqrt{n}x-\sqrt{n}\mu)^2}\sqrt{n}$

So the $\overline{X}$ is normally distributed with mean μ and variance σ^2/n.

$(n-1)S^2/\sigma^2=\frac{1}{\sigma^2}\sum_{k=2}^{n}Y_k^2=\sum_{k=2}^{n}(\frac{Y_k}{\sigma})^2$ is a chi-squared distribution with $n-1$ freedom.

Independence of Sample Mean and Sample Variance in more General Form