7-16 (790x1010x24b jpeg)

Basic inequality

For any $y_1 \dots y_n$, and any $\mu \in \mathbb{R}$

$$\boxed{\frac{1}{n}\sum_{j=1}^{n}(y_j-\bar{y})^2 = \frac{1}{n}\sum_{j=1}^{n}(y_j-\mu)^2 - (\bar{y}-\mu)^2}$$

Proof

Let X be a discrete r.v with frequency probabilities

x	$y_1-\mu$	$y_2-\mu$	$\dots$	$y_n-\mu$
$P(X=x)$	$\frac{1}{n}$	$\frac{1}{n}$		$\frac{1}{n}$

Note that μ is not the mean of X

Then

$$E[X] = (y_1-\mu)\frac{1}{n} + (y_2-\mu)\frac{1}{n} + \dots + (y_n-\mu)\frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n}(y_i-\mu) = \bar{y}-\mu$$

$$E[X^2] = (y_1-\mu)^2\frac{1}{n} + (y_2-\mu)^2\frac{1}{n} + \dots + (y_n-\mu)^2\frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n}(y_i-\mu)^2$$

$$E[(X-EX)^2] = (y_1-\mu-(\bar{y}-\mu))^2\frac{1}{n} + \dots + (y_n-\mu-(\bar{y}-\mu))^2\frac{1}{n}$$

$$= (y_1-\bar{y})^2\frac{1}{n} + \dots + (y_n-\bar{y})^2\frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2$$

Since $Var(X) = E[X^2] - (E[X])^2$

$$\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i-\mu)^2 - (\bar{y}-\mu)^2$$

Basic estimation problems

We want to study a population, having an unknown population mean μ and an unknown population variance σ^2. To estimate μ and σ^2, we take a random sample $y_1 \ldots y_n$,

(1) We estimate the population mean, using the sample mean $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$

(2) If we know the population mean μ, we estimate the population variance, using $\frac{1}{n} \sum_{i=1}^{n} (y_i - \mu)^2$

(3) If we do not know the population mean μ, we estimate the population variance, using the sample variance

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$$

To be able to do inferences, we need to know the distribution of $\bar{Y}$, $\frac{1}{n} \sum_{i=1}^{n} (Y_i - \mu)^2$ and $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$

If $Y_1 \ldots Y_n$ are iid r.v's from a $N(\mu, \sigma^2)$ distribution, we can find the distribution of $\bar{Y}$, $\frac{1}{n} \sum_{i=1}^{n} (Y_i - \mu)^2$ and S^2

Multiplying the basic inequality by $\frac{1}{\sigma^2}$, we get

$$\frac{1}{\sigma^2}\sum_{j=1}^{n}(y_j-\bar{y})^2 = \frac{1}{\sigma^2}\sum_{j=1}^{n}(y_j-\mu)^2 - \frac{n(\bar{y}-\mu)^2}{\sigma^2}$$

or
$$\sum_{j=1}^{n}\left(\frac{y_j-\mu}{\sigma}\right)^2 = \frac{1}{\sigma^2}\sum_{j=1}^{n}(y_j-\bar{y})^2 + \left(\frac{\sqrt{n}(\bar{y}-\mu)}{\sigma}\right)^2$$

We will prove that

(1) $\displaystyle\sum_{j=1}^{n}\left(\frac{y_j-\mu}{\sigma}\right)^2 \overset{d}{\sim} \chi^2(n)$

(2) $\displaystyle\frac{1}{\sigma^2}\sum_{j=1}^{n}(y_j-\bar{y})^2 \overset{d}{\sim} \chi^2(n-1)$

(3) $\displaystyle\left(\frac{\sqrt{n}(\bar{y}-\mu)}{\sigma}\right)^2 \overset{d}{\sim} \chi^2(1)$

or
$$\frac{1}{n}\sum_{j=1}^{n}(y_j-\mu)^2 \overset{d}{\sim} \frac{\sigma^2}{n}\chi^2(n)$$

$$s^2 = \frac{1}{n-1}\sum_{j=1}^{n}(y_j-\bar{y})^2 \overset{d}{\sim} \frac{\sigma^2}{n-1}\chi^2(n)$$

$$\bar{y} \overset{d}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$$

7-19 (659x900x24b jpeg)

Lemma

Let Y_1 and let Y_2 be two independent r.v's.

Let $Y_1 \overset{d}{\sim} \chi^2(\nu_1)$ and $Y_1 + Y_2 \overset{d}{\sim} \chi^2(\nu_1 + \nu_2)$

Then, $Y_2 \sim \chi^2(\nu_1 + \nu_2)$

Proof Since Y_1 and Y_2 are independent r.v's

$$E[e^{t(Y_1+Y_2)}] = E[e^{tY_1}]\,E[e^{tY_2}]$$

So, $$E[e^{tY_2}] = \frac{E[e^{t(Y_1+Y_2)}]}{E[e^{tY_1}]}$$

If $Y_1 \overset{d}{\sim} \chi^2(\nu_1)$, $E[e^{tY_1}] = (1-2t)^{-\frac{\nu_1}{2}}$

Similarly, $E[e^{t(Y_1+Y_2)}] = (1-2t)^{-\frac{(\nu_1+\nu_2)}{2}}$

So. $$E[e^{tY_2}] = \frac{(1-2t)^{-\frac{(\nu_1+\nu_2)}{2}}}{(1-2t)^{-\frac{\nu_1}{2}}} = (1-2t)^{-\frac{\nu_2}{2}}$$

which is the mgf of a $\chi^2(\nu_2)$

Hence, $Y_2 \overset{d}{\sim} \chi^2(\nu_2)$

Theorem

Let $Y_1, \ldots, Y_n$ be a random sample from a normal distribution with mean μ and variance σ^2. Then

(1) $\bar{Y}$ and S^2 are independent r.v.'s

(2) $\frac{n-1}{\sigma^2} S^2 \overset{d}{\sim} \chi^2(n-1)$

Proof Since $\bar{Y}$ and $Y_i - \bar{Y}$ are linear combinations of independent normal r.v.'s.

$\bar{Y}$ and $Y_i - \bar{Y}$ are independent iff $Cov(\bar{Y}, Y_i - \bar{Y}) = 0$

Now

$Cov(\bar{Y}, Y_i - \bar{Y}) = Cov(\bar{Y}, Y_i) - Cov(\bar{Y}, \bar{Y})$

Since $Cov(\bar{Y}, Y_i) = Cov(\bar{Y}, Y_j)$ $\qquad Cov(\bar{Y}, Y_i) = \frac{1}{n}\sum_{j=1}^{n} Cov(\bar{Y}, Y_j)$

$= Cov(\bar{Y}, \bar{Y})$, So, $Cov(\bar{Y}, Y_i - \bar{Y}) = 0$.

Since, for each $i = 1, \ldots n$, $\bar{Y}$ and $Y_i - \bar{Y}$ are independent r.v's.

$\bar{Y}$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ are independent r.v's.

Now, we use the Lemma $\underbrace{\sum_{j=1}^{n}\left(\frac{Y_j - \mu}{\sigma}\right)^2}_{\chi^2(n)} = \frac{1}{\sigma^2}\sum_{j=1}^{n}(Y_j - \bar{Y})^2 + \underbrace{\left(\frac{\sqrt{n}(\bar{Y} - \mu)}{\sigma}\right)^2}_{\chi^2(1)}$

and $\frac{1}{\sigma^2}\sum_{j=1}^{n}(Y_j - \bar{Y})^2$ and $\left(\frac{\sqrt{n}(\bar{Y} - \mu)}{\sigma}\right)^2$ are independent r.v's.