Development statistics 06

S06 Sample Distribution

Fujikawa, Kiyoshi Nagoya University, GSID

Distribution of Sample Mean

Population =Normal Distribution

$$N(\mu, \sigma^2)$$

Distribution of sample mean
Variance becomes smaller then that of the population.

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

Sample Distribution

Mean of sample mean

$$E(\overline{X}) = E(\frac{1}{n}\sum_{i}x_{i}) = \frac{1}{n}E(\sum_{i}x_{i})$$
$$= \frac{1}{n}E(\sum_{i}\mu) = \frac{n\mu}{n} = \mu$$

Sample Distribution

Variance of sample mean

$$Var(\overline{X}) = Var(\frac{1}{n}\sum_{i}x_{i})$$
$$= \frac{1}{n^{2}}Var(\sum_{i}x_{i})$$
$$= \frac{1}{n}Var(x_{i}) = \frac{1}{n}\sigma^{2}$$

Sample Distribution

Standardization of Normal distribution

Population =Normal Distribution

 $N(\mu, \sigma^2)$

Distribution of sample mean

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n}) \longrightarrow z = \frac{X - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0, 1)$$

Sample Distribution

Sample Variance

Maximum likelihood variance(最尤分散)
 Large sample theory
 Consistency (一致性)
 $s^2 = \frac{\sum_i (x_i - \bar{x})^2}{n}$ Unbiased variance
 small sample theory
 $\sum_i (x_i - \bar{x})^2$

●Unbiased (不偏性)

Basic knowledge

$$Var(x_i) = \sigma^2$$

= $E(x_i - \mu)^2$
= $E(x_i^2) - 2\mu E(x_i) + \mu^2$
= $E(x_i^2) - \mu^2$
 $E(x_i^2) = \sigma^2 + \mu^2$

Sample Distribution

mean of sample variance

$$E\sum_{i} (x_{i} - \overline{X})^{2}$$

$$= E\sum_{i} (x_{i}^{2}) - 2E(\overline{X}\sum_{i} x_{i}) + nE(\overline{X}^{2})$$

$$= E\sum_{i} (x_{i}^{2}) - 2nE(\overline{X}^{2}) + nE(\overline{X}^{2})$$

$$= E\sum_{i} (x_{i}^{2}) - nE(\overline{X}^{2})$$

$$= n(\sigma^{2} + \mu^{2}) - n(\frac{1}{n}\sigma^{2} + \mu^{2})$$

$$= (n-1)\sigma^{2}$$

Sample Distribution

8

Unbiased variance

$$E\frac{1}{n-1}\sum (x_i - \overline{X})^2 = \sigma^2$$

Sample Distribution

Chi-square distribution

The population is N(0,1)
Take sample whose size is n

$$x_i \sim N(0,1)$$

 The squared sum of *n* samples follows Chi-square distribution with degree of freedom *n* :χ² (*n*)

$$C = \sum_{i=1}^{n} x_i^2 \sim \chi^2(n)$$

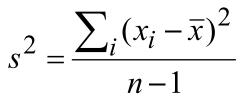
Sample Distribution

10

Distribution of sample variance

The population is N(μ,σ²)
Take samples whose size is n

Unbiased sample Variance



Chi square distribution

$$C = (n-1)\frac{s^2}{\sigma^2} \sim \chi^2(n-1)$$

Sample Distribution

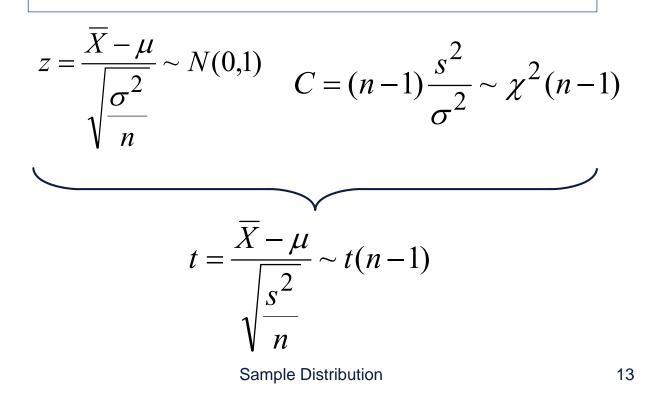
11

T distribution

 $z \sim N(0,1)$ $C \sim \chi^2(n)$ $t = \frac{Z}{\sqrt{C}} \sim t(n)$

Sample Distribution

Practically speaking



Binomial distribution

- Probability of "success or failure"
 - 1 (success) with probability p
 - 0 (failure) with probability 1-p
- Trial times : n
- Probability of total points "x"

 $P(x) =_{n} C_{x} p^{x} (1-p)^{n-x}$

Mean & variance of Binomial Distribution

Binomial distribution
 = total of Bernoulli distribution

Mean

$$\mu = 1 * p + 1 * p + \dots = np$$

Variance

$$\sigma^2 = 1^2 p + 1^2 p + \dots = np(1-p)$$

Sample Distribution

15

Distribution of proportion

Distribution of ratio

$$\frac{x}{n} = \hat{p} \sim N(p, p(1-p)/n)$$

$$z = \frac{\hat{p} - p}{\sqrt{p(1 - p)/n}} \sim N(0, 1)$$

If the sample size is large enough

$$z = \frac{\hat{p} - p}{\sqrt{\hat{p}(1 - \hat{p}) / n}} \sim N(0, 1)$$

Sample Distribution