

Lecture 20 Confidence Interval For The Variance

BIO210 Biostatistics

Xi Chen

Spring, 2022

School of Life Sciences

Southern University of Science and Technology



南方科技大学生命科学学院
SUSTech · SCHOOL OF
LIFE SCIENCES

Interval Estimation For The Variance

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \text{Not normally distributed!}$$

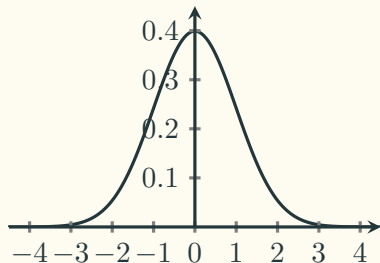
Sampling Distribution of The Sample Variance

$mean = 0$

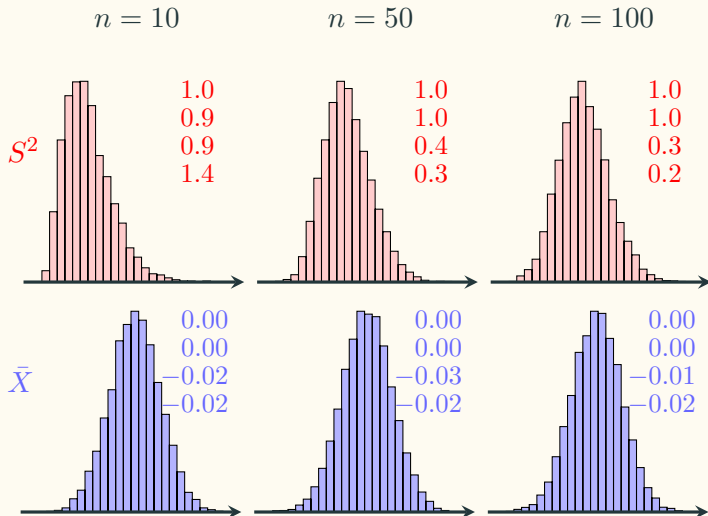
$median = 0$

$skew = 0$

$kurtosis = 0$

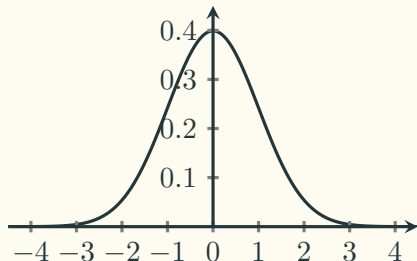


Population: $\mathcal{N}(0, 1)$



Chi-square Distribution

$$Q_1, Q_2, Q_3, \dots, Q_n \sim \mathcal{N}(0, 1)$$



$$G_1 = Q_1^2$$

$$G_1 \sim \chi_1^2$$

$$G_2 = Q_1^2 + Q_2^2$$

$$G_2 \sim \chi_2^2$$

$$G_3 = Q_1^2 + Q_2^2 + Q_3^2$$

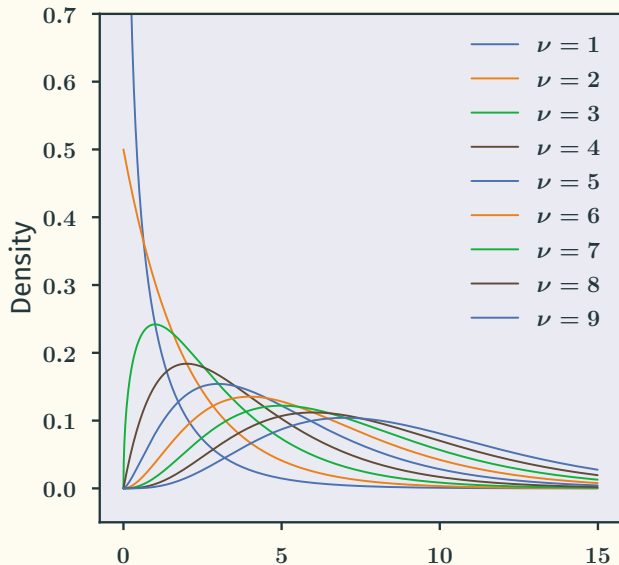
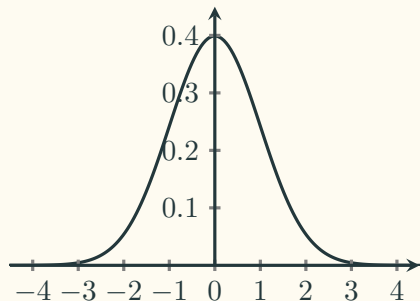
$$G_3 \sim \chi_3^2$$

$$\vdots$$
$$\vdots$$

$$G_n = Q_1^2 + Q_2^2 + Q_3^2 + \dots + Q_n^2$$

$$G_n \sim \chi_n^2$$

Chi-square Distributions of Different ν



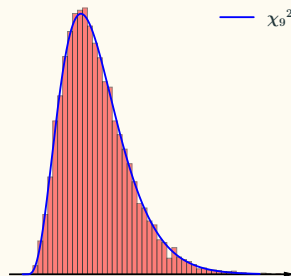
Sampling Distribution of The Sample Variance

Let X_1, X_2, \dots, X_n be a sample independently drawn from a normal distribution $\mathcal{N}(\mu, \sigma)$, and we define the sample mean and variance as:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Then we have:

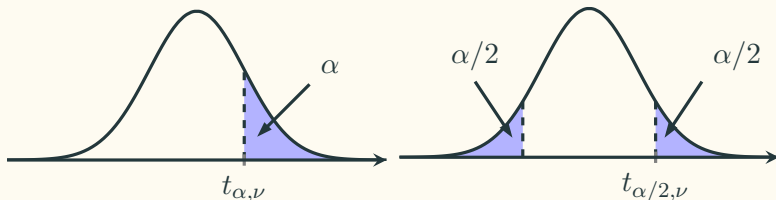
$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$



Confidence Interval For The Mean

Recall: construct $(1 - \alpha) \times 100\%$ confidence interval for the mean.

- One-sided (lower bound): $P\left(\mu \geq \bar{X} - t_{\alpha,\nu} \cdot \frac{s}{\sqrt{n}}\right) = 1 - \alpha$
- One-sided (upper bound): $P\left(\mu \leq \bar{X} + t_{\alpha,\nu} \cdot \frac{s}{\sqrt{n}}\right) = 1 - \alpha$
- Two-sided: $P\left(\bar{X} - t_{\alpha/2,\nu} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\alpha/2,\nu} \cdot \frac{s}{\sqrt{n}}\right) = 1 - \alpha$



Confidence Interval For Variance

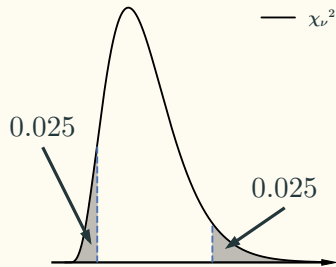
Goal: for a population with unknown variance σ^2 , find a and b , such that $P(a \leq \sigma^2 \leq b) = 0.95$

$$P(\chi_{0.975,\nu}^2 \leq \chi^2 \leq \chi_{0.025,\nu}^2) = 0.95$$

$$P\left[\chi_{0.975,\nu}^2 \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi_{0.025,\nu}^2\right] = 0.95$$

$$P\left[\frac{1}{\chi_{0.025,\nu}^2} \leq \frac{\sigma^2}{(n-1)S^2} \leq \frac{1}{\chi_{0.975,\nu}^2}\right] = 0.95$$

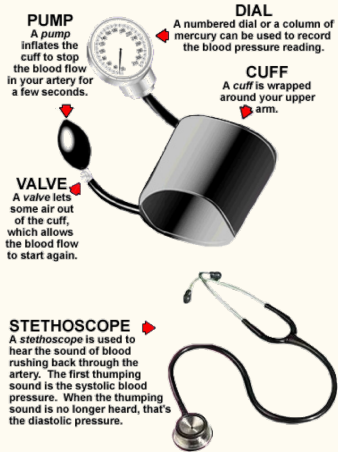
$$P\left(\frac{n-1}{\chi_{0.025,\nu}^2} \cdot S^2 \leq \sigma^2 \leq \frac{n-1}{\chi_{0.975,\nu}^2} \cdot S^2\right) = 0.95$$



95% CI for σ^2 :

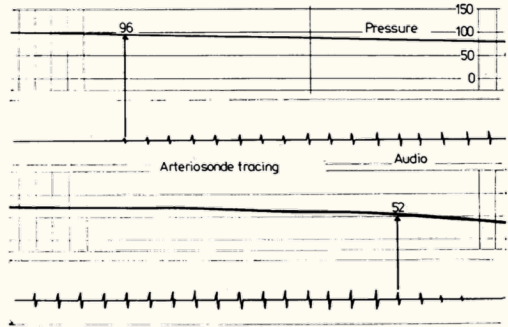
$$\left(\frac{n-1}{\chi_{0.025,\nu}^2} \cdot S^2, \frac{n-1}{\chi_{0.975,\nu}^2} \cdot S^2\right)$$

Sphygmomanometer



Confidence interval for variance

Hypertension: An Arteriosonde machine “prints” blood-pressure readings on a tape so that the measurement can be read rather than heard. A major argument for using such a machine is that the variability of measurements obtained by different observers on the same person will be lower than with a standard blood-pressure cuff.



Confidence interval for variance

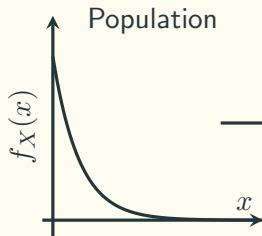
Persion (i)	Observer #1	Observer #2	Difference (d_i)
1	194	200	-6
2	126	123	3
3	130	128	2
4	98	101	-3
5	136	135	1
6	145	145	0
7	110	111	-1
8	108	107	1
9	102	99	3
10	126	128	-2

Does it make sense to construct a 95% CI for the variance?

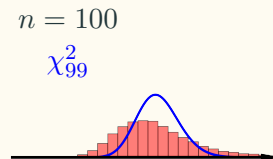
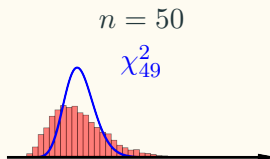
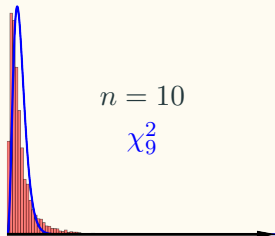
95% CI for σ^2 :

$$\left(\frac{n-1}{\chi_{0.025,\nu}^2} \cdot S^2, \frac{n-1}{\chi_{0.975,\nu}^2} \cdot S^2 \right)$$

Sampling Distribution of The Sample Variance



Sample (n): $\frac{(n-1)S^2}{\sigma^2}$



Conditions For Valid Confidence Intervals For The Variance

1. Random Samples
2. Original population distribution must be normal
3. Independence ($n < 10\%$ population size)