

# Lecture 20 Confidence Interval For The Variance

BIO210 Biostatistics

---

Xi Chen

Spring, 2024

School of Life Sciences

Southern University of Science and Technology



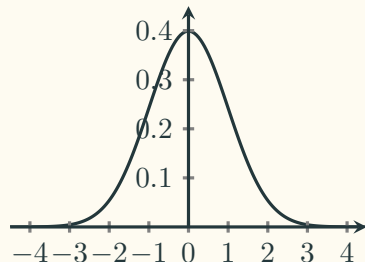
南方科技大学生命科学学院  
SUSTech · SCHOOL OF  
**LIFE SCIENCES**

## Interval Estimation For The Variance

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

# Chi-squared Distribution

$$Z_1, Z_2, Z_3, \dots, Z_n \sim \mathcal{N}(0, 1)$$



$$U_1 = Z_1^2$$

$$U_1 \sim \chi^2(1)$$

$$U_2 = Z_1^2 + Z_2^2$$

$$U_2 \sim \chi^2(2)$$

$$U_3 = Z_1^2 + Z_2^2 + Z_3^2$$

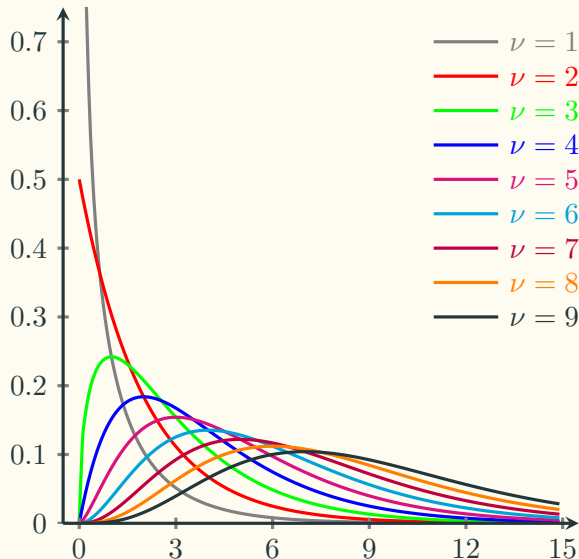
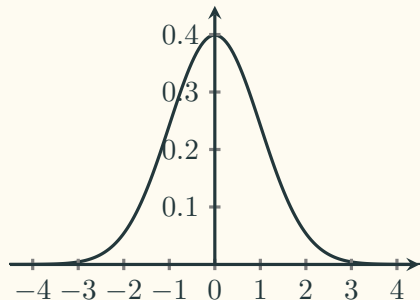
$$U_3 \sim \chi^2(3)$$

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

$$U_n = Z_1^2 + Z_2^2 + Z_3^2 + \dots + Z_n^2$$

$$U_n \sim \chi^2(n)$$

## Chi-squared Distributions of Different $\nu$



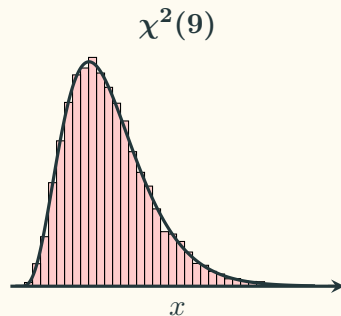
# Sampling Distribution of The Sample Variance

Let  $X_1, X_2, \dots, X_n$  be a sample independently drawn from a normal distribution  $\mathcal{N}(\mu, \sigma^2)$ , 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:

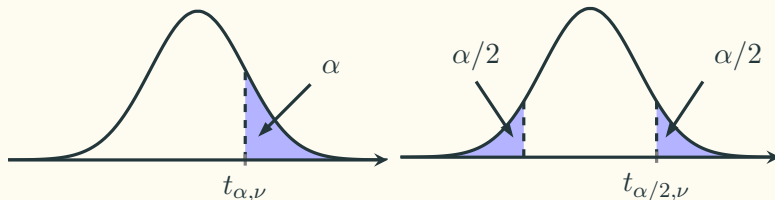
$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$



# Confidence Interval For The Mean

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

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



## Confidence Interval For Variance

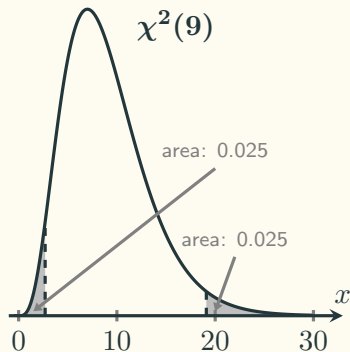
**Goal:** for a population with unknown variance  $\sigma^2$ , find  $a$  and  $b$ , such that  $\mathbb{P}(a \leq \sigma^2 \leq b) = 0.95$

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

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

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

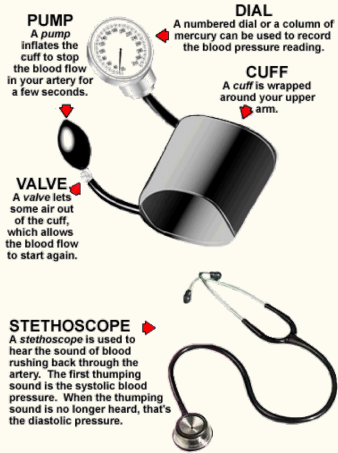
$$\mathbb{P}\left(\frac{n-1}{\chi_{0.025, \nu}^2} \cdot S^2 \leq \sigma^2 \leq \frac{n-1}{\chi_{0.975, \nu}^2} \cdot S^2\right) = 0.95$$



95% CI for  $\sigma^2$  :

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

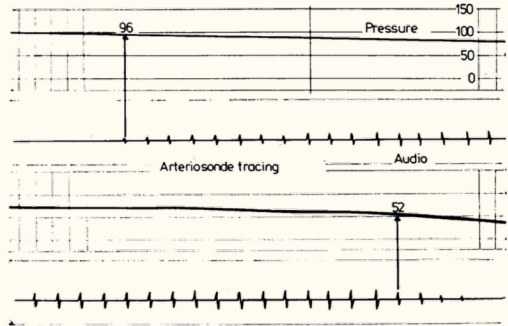
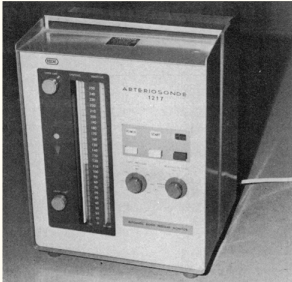
# Sphygmomanometer





# 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

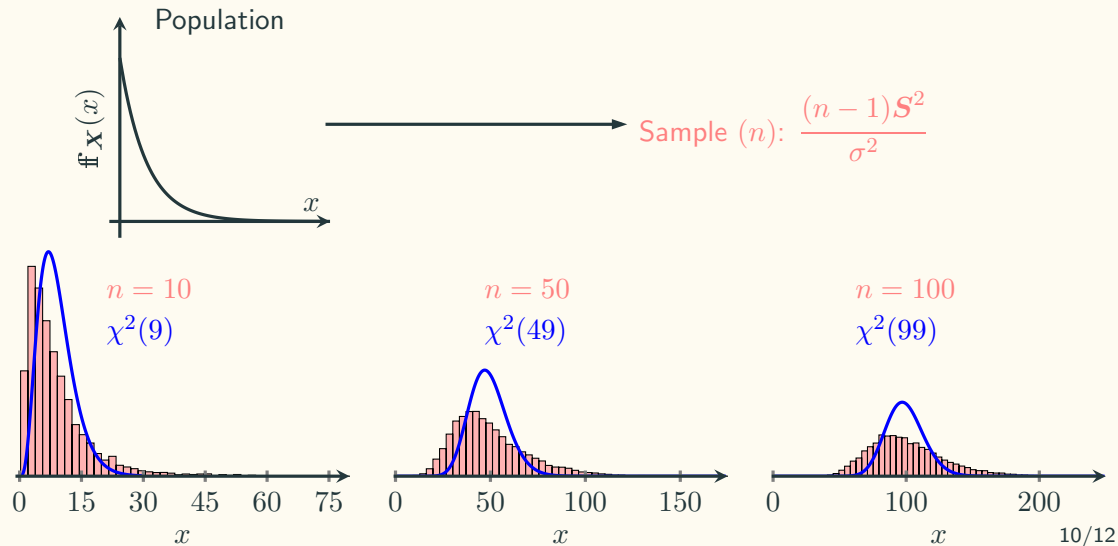
Persion (i)	Observer #1	Observer #2	Difference ( $d_i$ )
1	194	200	-6
2	126	123	3
3	130	128	2
4	98	101	-3
5	136	135	1
6	145	145	0
7	110	111	-1
8	108	107	1
9	102	99	3
10	126	128	-2

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

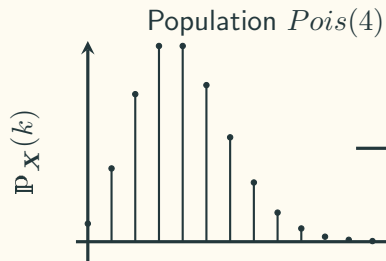
95% CI for  $\sigma^2$  :

$$\left[ \frac{n-1}{\chi_{0.025,\nu}^2} \cdot s^2, \frac{n-1}{\chi_{0.975,\nu}^2} \cdot s^2 \right]$$

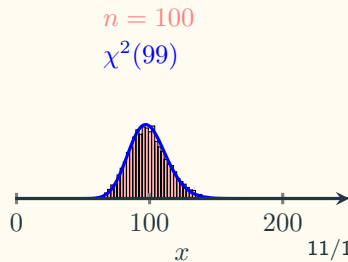
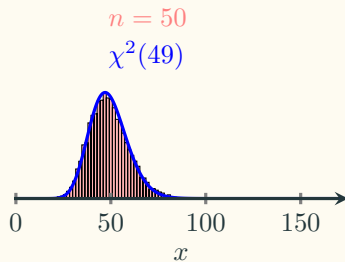
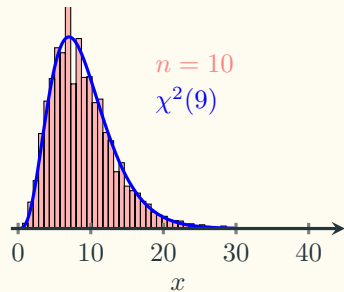
# Sampling Distribution of The Sample Variance



# 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. Independence ( $n < 10\%$  population size)
3. Original population distribution must be normal