STA 130A: Mathematical Statistics: Brief Course

Lecture 23: Sampling Distributions of Means and Variances

Dogyoon Song

Fall 2026, UC Davis

Announcements

Homework 6 is now due Sun, Nov 22, 11:59 PM PT

Homework 7 is posted

- Due Thu, Dec 3, 11:59 PM PT
- Please submit on time and follow the submission instructions
- Review the problems early so you have time to ask questions

Agenda

Last time

- A parameter is fixed; an estimator varies across random samples.
- Different samples can give different estimates of the same quantity.

Today

- Sampling distributions and standard errors
- Sample variance and the divisor $n - 1$
- Chi-square distributions and exact results for a normal sample

Goal: Understand how much our sample summaries can vary.

Recap: From a random sample to an estimate

Setting: Draw a sample from a distribution with unknown parameter θ. Our goal is to estimate θ from the data.

$$\underbrace{X_1, \ldots, X_n}_{\text{random sample}} \xrightarrow{\text{observe}} \underbrace{x_1, \ldots, x_n}_{\text{observed data}}$$

Definition (Statistic)

A **statistic** $T(X_1, \ldots, X_n)$ is computable from the data without knowing θ.
The **rule** T **is fixed**; its output is **random across samples**.

	Meaning	Mean-weight example
Parameter	Unknown population target (fixed)	μ
Estimator	Statistic estimating θ (random)	$\widehat{\mu} = \bar{X}$
Estimate	Observed value (deterministic)	$\bar{x} = 501.2$ grams

Distinction: The population determines the parameter's value. We choose the estimator.

Motivation: The same rule, different samples

Suppose we repeatedly take samples of four packages from the same machine, and apply the **same averaging rule** each time.

Example (Three hypothetical samples)

Sample	Observed weights (grams)	Observed mean
A	$498, 500, 502, 504$	501
B	$496, 498, 500, 502$	499
C	$499, 501, 503, 505$	502

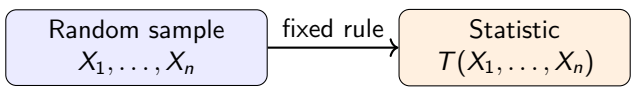
The population mean μ stays fixed; the statistic $\bar{X}$ varies from sample to sample.

Question: What distribution describes these possible sample means $\bar{X}$?

Sampling distribution of a statistic

Definition (Sampling distribution)

The **sampling distribution** is the probability distribution of $T(X_1, \ldots, X_n)$ under the sampling model.



- Imagine repeating the *entire sampling procedure*, not reusing the same observed data
- A sampling distribution concerns the statistic, not the individual observations
- Its shape and spread depend on the population model and the sample size

Example: Mean and variance of the sample mean

For i.i.d. observations with mean μ and variance σ^2,

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

By linearity,

$$\mathbb{E}[\bar{X}] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X_i] = \mu.$$

By independence and variance scaling,

$$\mathrm{Var}(\bar{X}) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{\sigma^2}{n}.$$

Statistical meaning: $\bar{X}$ is centered at its target μ, and its variance decreases as n grows.

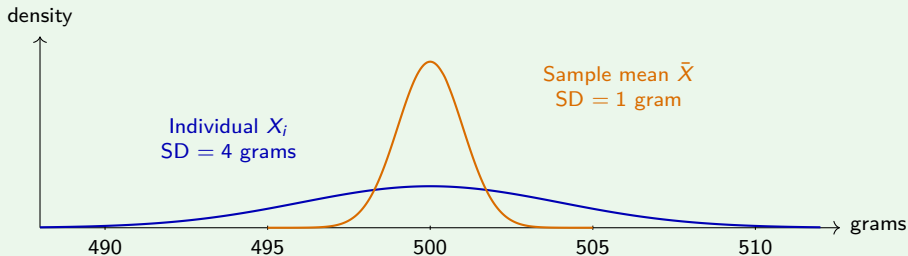
- *Note:* These moment identities do not require a normal population.

Example: Individual weights versus sample means

Example

For illustration, suppose $X_i \sim N(500, 4^2)$ independently and $n = 16$.

Then $\bar{X} \sim N(500, 1^2)$, exactly.



Same center, different spread. Averaging reduces sampling variability of $\bar{X}$.

- It does not reduce the variability of individual package weights X_i.

Standard deviation versus standard error

Population standard deviation:

$$\mathrm{SD}(X_i) = \sigma.$$

It measures the spread of individual observations.

Standard error of the sample mean:

$$\mathrm{SE}(\bar{X}) = \mathrm{SD}(\bar{X}) = \frac{\sigma}{\sqrt{n}}.$$

It measures the spread of the estimator across repeated samples.

Remarks:

- A standard error is not the realized error $\bar{x} - \mu$.
- Population SD and standard error have the same units, but describe different variability.

Example: What does a larger sample buy?

Example (Package weights)

Suppose the population standard deviation is $\sigma = 4$ grams.

Sample size	Individual SD	Standard error of $\bar{X}$
$n = 16$	4 grams	$4/\sqrt{16} = 1$ gram
$n = 64$	4 grams	$4/\sqrt{64} = 0.5$ gram

Question: How much more data is needed to halve the standard error?

Answer: Four times as many observations, because standard error decreases as $1/\sqrt{n}$. (But the standard deviation of individual sample remains the same.)

Caution: More observations reduce sampling variability, not systematic sampling bias.

Caveat: Exact normal sampling versus the CLT

Normal population: If $X_i \sim N(\mu, \sigma^2)$ independently, then

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \qquad \text{for every } n.$$

This is an exact distributional result, not an asymptotic approximation.

General i.i.d. population: The CLT gives, for large n,

$$P(\bar{X} \le c) \approx \Phi\left(\frac{c-\mu}{\sigma/\sqrt{n}}\right), \qquad \sigma > 0.$$

The quality of this approximation depends on the population and sample size.

In either case: $\mathbb{E}[\bar{X}] = \mu$ and $\mathrm{Var}(\bar{X}) = \sigma^2/n$. The individual observations do not become normal merely because n grows.

Pop-up quiz: What changes across samples?

Suppose we repeatedly draw samples of the same size from one unchanged population.

Question: Which statement is correct?

A) The population mean μ changes, but $\bar{X}$ is fixed.

B) Both μ and $\bar{X}$ must change.

C) The population mean μ is fixed, but $\bar{X}$ can change.

D) The standard error equals the observed error $\bar{x} - \mu$.

Answer: C. The estimator is random across samples; the population parameter stays fixed.

Follow-up: Once a sample has been observed, is its numerical mean $\bar{x}$ still an unknown outcome of that sampling experiment?

The sample variance and estimated standard error

The population variance σ^2 is usually unknown. For $n \geq 2$, define

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

- The **sample variance** S^2 estimates the population variance, and the **sample standard deviation** S estimates the population SD.
- After observing the data, compute s^2 and s using the same formulas.
- Variances have squared units; standard deviations have the original units.

Estimating the variability of the sample mean:

$$\underbrace{S/\sqrt{n}}_{\text{estimates } \mathrm{SE}(\bar{X})=\sigma/\sqrt{n}} \quad \text{and} \quad \underbrace{\frac{S^2}{n}}_{\text{estimates } \mathrm{Var}(\bar{X})=\sigma^2/n}.$$

Example: Compute and interpret three summaries

Example

(Sample mean) For the observed package weights $498, 500, 502, 504$ grams,

$$\bar{x} = \frac{498 + 500 + 502 + 504}{4} = 501 \text{ grams.}$$

(Sample standard deviation) The deviations from $\bar{x}$ are $-3, -1, 1, 3$, so

$$s^2 = \frac{9 + 1 + 1 + 9}{4 - 1} = \frac{20}{3} \text{ grams}^2, \qquad s \approx 2.58 \text{ grams.}$$

(Estimated standard error of the mean)

$$\frac{s}{\sqrt{4}} = \sqrt{\frac{5}{3}} \approx 1.29 \text{ grams.}$$

- 501: our estimate of the population mean.
- 2.58: estimated spread of individual package weights.
- 1.29: estimated sampling spread of the mean, not its actual error.

Why divide by $n-1$?

We center the observations at the *sample* mean instead of the unknown population mean

$$\sum_{i=1}^{n}(X_i - \bar{X}) = 0.$$

Once $n-1$ centered deviations are specified, the last is determined by this constraint.

The exact expectation calculation gives

$$\mathbb{E}\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = (n-1)\sigma^2.$$

Consequently,

$$\mathbb{E}[S^2] = \sigma^2.$$

Meaning: The sample variance is correct *on average* across samples (we call this *unbiasedness*). It need not equal σ^2 for a particular sample.

Self-study: Why S^2 is unbiased

The sum-of-squares identity is

$$\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n}(X_i - \mu)^2 - n(\bar{X} - \mu)^2.$$

The cross terms cancel because $\sum_i (X_i - \bar{X}) = 0$.

Taking expectations and using $\mathbb{E}[\bar{X}] = \mu$,

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right] &= n\sigma^2 - n\mathrm{Var}(\bar{X}) \\ &= n\sigma^2 - n\frac{\sigma^2}{n} = (n-1)\sigma^2.\end{aligned}$$

Thus $\mathbb{E}[S^2] = \sigma^2$.

No normality assumption is needed for this expectation identity. The exact distribution of S^2 is a different question, described by the normal-sample theorem in this lecture.

Pop-up quiz: What does unbiasedness mean?

Let X_1, X_2 be i.i.d. $\text{Bernoulli}(1/2)$, so $\sigma^2 = 1/4$. For $n = 2$, the divisor $n - 1$ equals 1, so

$$S^2 = (X_1 - \bar{X})^2 + (X_2 - \bar{X})^2, \qquad \bar{X} = \frac{X_1 + X_2}{2}.$$

Question: Which statement is correct?

A. $S^2 = 1/4$ for every sample.

B. S^2 is biased because it never equals $1/4$.

C. S^2 is unbiased even though it never equals $1/4$.

D. S^2 is unbiased only if the population is normal.

Answer: C. The observations match with probability $1/2$, giving $S^2 = 0$. Otherwise, $S^2 = 1/2$.

$$\mathbb{E}[S^2] = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} = \sigma^2.$$

Takeaway: Unbiasedness concerns the average across repeated samples.

Motivation: Is the machine more variable than claimed?

Example (Checking consistency of package weights)

A machine's claimed population SD is 4 grams. A new random sample of 16 packages has a sample SD of 6 grams:

$$\text{Claim: } \sigma = 4 \qquad \text{Observed data: } n = 16, \quad s = 6.$$

Assume i.i.d. normal package weights; the population mean may be unknown.

Question: Could this much sample variation plausibly arise if the claim were correct?

$$\text{If } \sigma = 4, \qquad P(S \geq 6) = ?$$

Knowing $\mathbb{E}[S^2] = \sigma^2$ tells us the average across samples, but does not answer this question.

What we need: The sampling distribution of S^2, so we can judge the observed spread against the claim. We begin our discussion with the simpler normal-sample case.

Chi-square: A sum of squared standard normals

Definition (Chi-square distribution)

If $Z_1, \ldots, Z_\nu$ are independent $N(0,1)$ random variables, then

$$U = Z_1^2 + \cdots + Z_\nu^2 \sim \chi^2_\nu.$$

The positive integer ν is the number of **degrees of freedom**.

Example (Three packages: known mean 500, known SD 4)

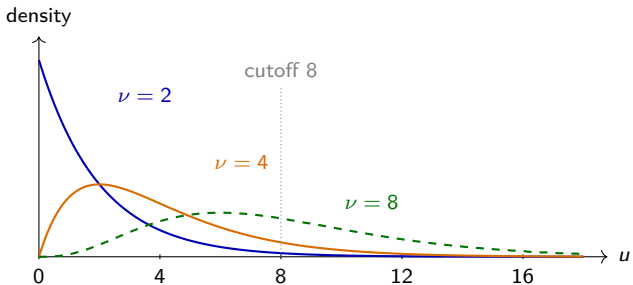
For independent $X_i \sim N(500, 4^2)$, let $Z_i = (X_i - 500)/4$. Then $U = Z_1^2 + Z_2^2 + Z_3^2 \sim \chi^2_3$. For observed weights $496, 500, 508$,

$$(z_1, z_2, z_3) = (-1, 0, 2) \quad \Longrightarrow \quad u = 1 + 0 + 4 = 5.$$

The statistic U varies across samples; 5 is one observed value.

Intuition: Measure deviations in SD units, square them, and add. For S^2, we instead center at the sample mean; we will account for this next.

Chi-square: Shape and degrees of freedom



How to read these curves

- Squared terms make $U \geq 0$.
- More terms give a larger typical total; the curve becomes less skewed.
- Area to the right of a cutoff is the chance of exceeding it.

If $U \sim \chi^2_\nu$, then

$$\mathbb{E}[U] = \nu, \qquad \mathrm{Var}(U) = 2\nu.$$

Example: Exceeding 8 is less likely for $\nu = 2$ than for $\nu = 8$.

Message: Use the correct degrees of freedom to judge an observed total.

Normal-sample theorem: Variance and independence

Theorem (Exact results for a normal sample)

For i.i.d. $X_i \sim N(\mu, \sigma^2)$*, with* $n \geq 2$ *and* $\sigma > 0$*,*

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}, \qquad \bar{X} \text{ and } S^2 \text{ are independent.}$$

Meaning: Chi-square describes how S^2 varies relative to σ^2.

- Centering at $\bar{X}$ leaves $n-1$ degrees of freedom. The centered deviations are dependent; the result requires a theorem.
- Normal sampling is essential for these exact conclusions.

Example (Our inspection, if the SD claim is correct)

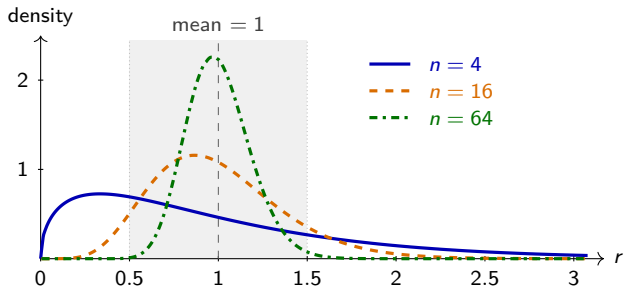
With $n = 16$, claimed $\sigma = 4$, and observed $s = 6$,

$$U = \frac{15S^2}{4^2} \sim \chi^2_{15}, \qquad u_{\text{obs}} = \frac{15 \cdot 6^2}{4^2} = 33.75.$$

Visualization: Accuracy of the sample variance

For i.i.d. normal samples, the theorem gives

$$R = \frac{S^2}{\sigma^2} = \frac{U}{\nu}, \qquad U \sim \chi^2_\nu, \qquad \nu = n - 1.$$



Gray band: S^2 is within 50% of σ^2.

Mean and shrinking spread

$$\mathbb{E}[R] = 1, \quad \mathrm{Var}(R) = \frac{2}{n-1}.$$

Error of at least 50%

n	$P(\lvert R - 1\rvert \geq 0.5)$
4	53.00%
16	15.31%
64	0.66%

Exact probabilities: Chi-square tails quantify errors in estimating σ^2.

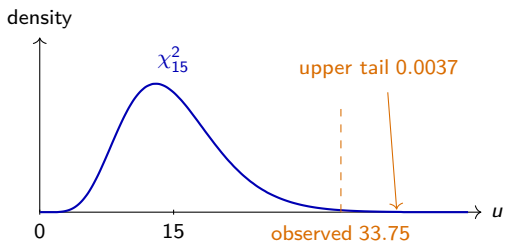
Concentration: $S^2/\sigma^2 \xrightarrow{P} 1$, so $S^2 \xrightarrow{P} \sigma^2$ as $n \to \infty$.

Example: Using the data to assess the SD claim

Example (Return to the sample of 16 packages)

We observed $s = 6$ grams. If the claimed $\sigma = 4$ normal model is correct,

$$P(S \geq 6) = P(U \geq 33.75) \approx 0.0037, \qquad U \sim \chi^2_{15}.$$



Probability calculation

Under the claim, only about **0.37%** of samples of 16 packages would have an SD of at least 6 grams.

The tail probability is supplied by a table or software.

Statistical conclusion: Under i.i.d. normal sampling, the observed spread casts doubt on the claimed SD of 4 grams.

Caution: 0.37% is not the probability that the SD claim is true.

Self-study: A probability involving the sample variance

Example (Package weights)

Suppose $X_1, \ldots, X_{16}$ are independent $N(\mu, 4^2)$ package weights.

Question: What is $P(S > 5)$, the chance that the sample SD exceeds 5 grams?

$$U = \frac{15S^2}{16} \sim \chi^2_{15}.$$

Since $S \geq 0$,

$$P(S > 5) = P(S^2 > 25) = P\left(U > \frac{15 \cdot 25}{16}\right) = P(U > 23.4375) \approx 0.0753.$$

The final chi-square tail probability is supplied by a table or software.

Interpretation: Sample SDs vary even when the population SD stays fixed.

Workflow: Transform the variance cutoff, then use the correct degrees of freedom.

(Optional) Which results need normality?

Result	Sampling model
$\mathbb{E}[\bar{X}] = \mu, \quad \mathrm{Var}(\bar{X}) = \sigma^2/n$	General i.i.d. sample
$\mathbb{E}[S^2] = \sigma^2$	General i.i.d. sample
$\bar{X} \sim N(\mu, \sigma^2/n)$, exactly	I.i.d. normal sample
$(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$, exactly	I.i.d. normal sample
$\bar{X}$ and S^2 are independent	I.i.d. normal sample

Quick check: If the observations are exponential, is S^2 still unbiased?

Answer: Yes. The unbiasedness identity does not require a normal population.

Next question: What changes when σ is unknown?

For i.i.d. $N(\mu, \sigma^2)$ observations with $n \geq 2$ and $\sigma > 0$,

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

But σ is usually unknown; replacing it by S gives

$$\frac{\bar{X} - \mu}{S/\sqrt{n}}.$$

- The denominator is now random too.
- This ratio is not exactly standard normal under normal sampling.
- The chi-square law and mean–variance independence let us identify its distribution.

Next lecture: Student's t and confidence intervals for a mean.

Wrap-up

Sampling distributions describe how statistics vary across samples.

$$\mathbb{E}[\bar{X}] = \mu, \qquad \mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}, \qquad \mathrm{SE}(\bar{X}) = \frac{\sigma}{\sqrt{n}}.$$

- S^2 estimates population variance; $S/\sqrt{n}$ estimates the SE of the mean.
- Unbiasedness describes an average across samples, not an exact match in one sample.
- Under normal sampling, scaled S^2 is chi-square and is independent of $\bar{X}$.

Suggested reading: [Ric07, Ch. 6].

Next lecture: Student's t and confidence intervals for a mean.

References


John A Rice.
Mathematical statistics and data analysis, volume 9.
Thomson brooks/cole, Belmont, CA, 2007.