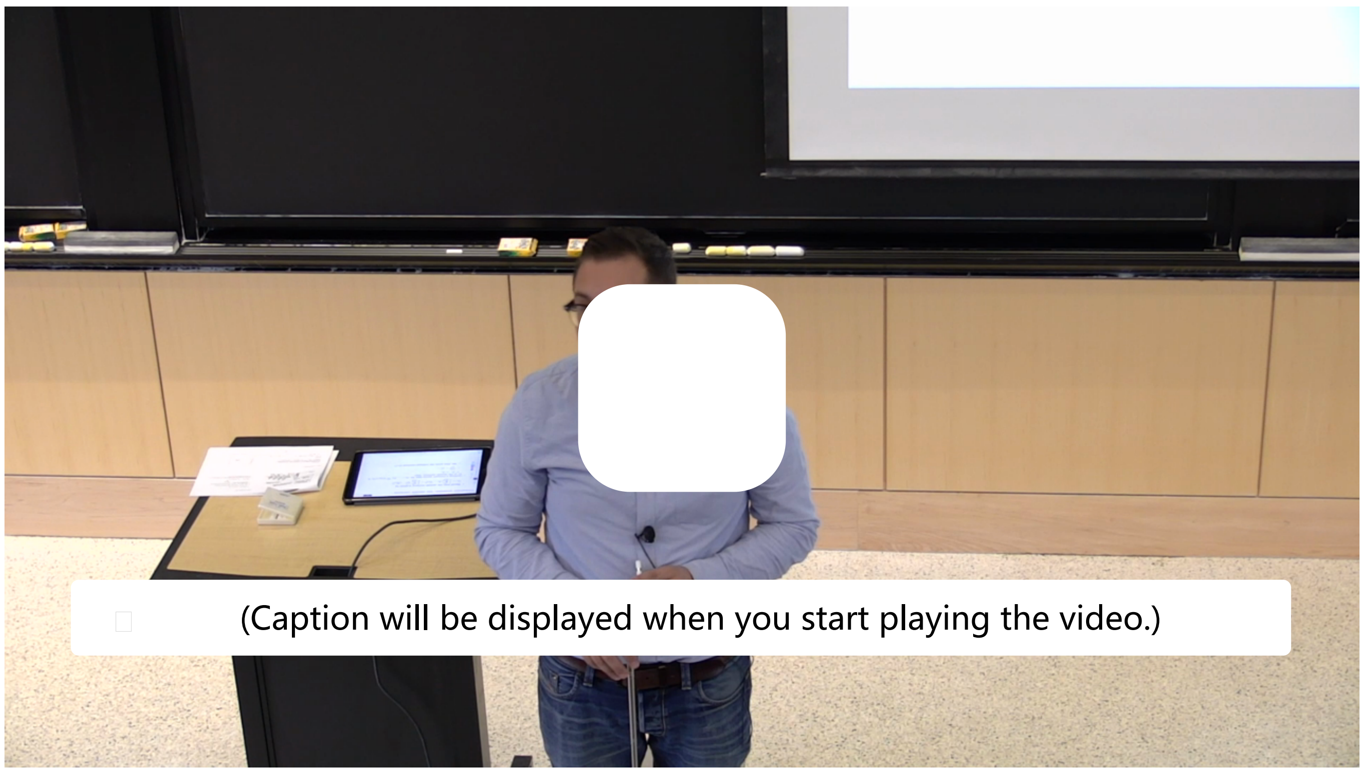
8. Sample Variance and Sample Mean of IID Gaussians: Cochran's Theorem

Cochran's Theorem: Independence of Gaussian Sample Variance and Sample Mean



Start of transcript. Skip to the end.

So I promised you that we'll talk about this being the law of the variance.

And so this is probably the most important example.

And there's actually two important results that are on this slide.

The first one says that, if you actually look at the sample variance--

so what is the sample variance?

视频
下载视频文件

字幕
下载 SubRip (.srt) file
下载 Text (.txt) file

A Special Case of Cochran's Theorem I

3/3 points (graded)

Cochran's theorem states that if $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, then the sample variance

$$S_n := \frac{1}{n}\left(\sum_{i=1}^{n} X_i^2\right) - \left(\overline{X}_n\right)^2$$

satisfies:

- $\overline{X}_n$ is independent of S_n, and
- $\frac{nS_n}{\sigma^2} \sim \chi^2_{n-1}$.

In this problem, you will verify that Cochran's theorem holds when $n = 2$. Let $X_1, X_2 \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$.

The expression S_2 can be written in the form A^2 where A is a polynomial in X_1 and X_2.

What is A^2?

Type **X_1** for X_1 and **X_2** for X_2.

$A^2 =$ ((X_1 - X_2)/2)^2 Answer: (X_1 - X_2)^2/4

$\left(\frac{X_1-X_2}{2}\right)^2$

STANDARD NOTATION

The expression A from the previous question is a random variable, and moreover is distributed as $\mathcal{N}\left(\mu^*, (\sigma^*)^2\right)$ for some μ^* and σ^* that can be expressed in terms of the original parameters μ and σ. (Note: A can have two forms, but both would have the same distribution by symmetry).

What is μ^* expressed in terms of μ and σ?

$\mu^* =$ 0 Answer: 0.0

0

What is $(\sigma^*)^2$ expressed in terms of μ and σ?

$(\sigma^*)^2 =$ sigma^2/2 Answer: sigma^2/2

$\frac{\sigma^2}{2}$

STANDARD NOTATION

Solution:

Observe that

$$S_n = \frac{X_1^2 + X_2^2}{2} - \left(\frac{X_1 + X_2}{2}\right)^2 = \frac{X_1^2}{4} + \frac{X_2^2}{4} - \frac{1}{2}X_1X_2 = \left(\frac{X_1 - X_2}{2}\right)^2.$$

Hence, we can take $A = \pm\frac{X_1-X_2}{2}$ (either choice has the same distribution, by symmetry). Next,

$$\mathbb{E}[A] = \mathbb{E}[X_1 - X_2] = \mu - \mu = 0,$$

and

$$\textbf{Var}(A) = \textbf{Var}\left(\frac{X_1 - X_2}{2}\right) = \frac{1}{4}(\textbf{Var}(X_1) + \textbf{Var}(X_2)) = \frac{\sigma^2}{2}.$$

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

A Special Case of Cochran's Theorem II

4/4 points (graded)

As above, let $X_1, X_2 \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$.

Recall the random variable A that you found in the previous problem in terms of X_1 and X_2.

Let $\overline{X}_2 = \dfrac{X_1 + X_2}{2}$, i.e. $\overline{X}_n$ when $n = 2$.

What is $\mathbb{E}[A\overline{X}_2]$?

0 Answer: 0.0

Using the answer to the previous question, which of the following are true? (Choose all that apply.)

- [x] A and $\overline{X}_2$ are independent.
- [ ] A and $\overline{X}_2$ are not independent.
- [ ] $A, \overline{X}_2 \sim \mathcal{N}(0, 2\sigma^2)$.
- [x] $A \sim \mathcal{N}(0, \sigma^2/2)$ and $\overline{X}_2 \sim \mathcal{N}(\mu, \sigma^2/2)$.

For some expression B in terms of σ^2, the random variable $BS_2 \sim \chi^2$. What is B?

$B =$ 2/sigma^2 Answer: 2/sigma^2

$\frac{2}{\sigma^2}$

STANDARD NOTATION

How many degrees of freedom does the χ^2 random variable BS_2 have?

1 Answer: 1

Solution:

Recall that $A = \frac{X_1 - X_2}{2}$ and $\overline{X}_2 = \frac{X_1 + X_2}{2}$. Hence,

$$\mathbb{E}[A\overline{X}_2] = \frac{1}{4}\mathbb{E}[(X_1 - X_2)(X_1 + X_2)] = \frac{1}{4}(\sigma^2 - \sigma^2) = 0.$$

A random vector $\mathbf{X} = (X^{(1)}, \ldots, X^{(d)})^T$ is a **Gaussian vector**, or **multivariate Gaussian or normal variable**, if any linear combination of its components is a (univariate) Gaussian variable or a constant (a "Gaussian" variable with zero variance), i.e., if $\alpha^T \mathbf{X}$ is (univariate) Gaussian or constant for any constant non-zero vector $\alpha \in \mathbb{R}^d$.

As jointly Gaussian variables (why is it that A and $\overline{X}_2$ are jointly Gaussian?) that are uncorrelated are also independent, A and $\overline{X}_2$ are independent. By the previous problem, we know $A \sim \mathcal{N}(0, \sigma^2/2)$. A quick calculation shows that $\overline{X}_2 \sim \mathcal{N}(\mu, \sigma^2/2)$. Hence, the first and last choices are correct in the multiple choice question.

Observe that

$$\boxed{\frac{2}{\sigma^2}S_2 = \frac{2}{\sigma^2}\left(\frac{X_1 - X_2}{2}\right)^2 = \left(\frac{X_1 - X_2}{\sqrt{2}\sigma}\right)^2},$$

and $\boxed{\frac{X_1 - X_2}{\sqrt{2}\sigma} \sim \mathcal{N}(0, 1)}$. By definition, $\boxed{\frac{2}{\sigma^2}S_2 \sim \chi_1^2}$.

Remark: The last question shows that $\frac{2}{\sigma^2}S_2 \sim \chi_1^2$, which verifies the second claim in Cochran's theorem for this special case. To show the first part of Cochran's theorem, that S_2 and $\overline{X}_2$ are independent, recall that we showed $A = \sqrt{S_2}$ is independent of $\overline{X}_2$. By a standard fact of probability, this also implies that $A^2 = S_2$ is independent of $\overline{X}_2$.

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

Concept Check: Cochran's Theorem and Unbiased Sample Variance

1/1 point (graded)

Let $X_1, \ldots, X_n$ be i.i.d. and distributed according to $\mathcal{N}(0, \sigma^2)$. Let $\overline{X}_n = \frac{X_1+X_2+\cdots+X_n}{n}$. What is the distribution of $\frac{(n-1)\widetilde{S}_n}{\sigma^2}$, where $\widetilde{S}_n$ is the the unbiased sample variance of $X_1, \ldots, X_n$:

$$\widetilde{S}_n = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2$$

Type **Cn** for chi-squared distribution with n degrees of freedom, **Cn1** for chi-squared distribution with $n-1$ degrees of freedom.

Cn1

Answer: Cn1 + 0*Cn

$Cn1$

STANDARD NOTATION

Solution:

By Cochran's theorem,

$$\frac{nS_n}{\sigma^2} \sim \chi^2_{n-1}$$
$$\iff \frac{(n-1)\,\widetilde{S}_n}{\sigma^2} \sim \chi^2_{n-1}$$

Remark: We will use the random variable $\frac{\widetilde{S}_n}{\sigma^2}$ in the upcoming videos in what is called the Student's T Test. The point of this problem was to show that $\frac{(n-1)\widetilde{S}_n}{\sigma^2}$ is a χ^2_{n-1} random variable, thereby showing that the distribution of $\frac{\widetilde{S}_n}{\sigma^2}$ is the distribution of a χ^2_{n-1} random variable scaled by $\frac{1}{n-1}$.

提交 你已经尝试了1次（总共可以尝试1次）

Answers are displayed within the problem

讨论

显示讨论

主题： Unit 4 Hypothesis testing:Lecture 13: Chi Squared Distribution, T-Test / 8. Sample Variance and Sample Mean of IID Gaussians: Cochran's Theorem