Cochran's theorem and T test

02 August 2020 14:05

1. Proof of Cochran's Theorem and T Test

Bookmark this page

Recitation problem statement

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \mathcal{N}(\mu, \sigma^2)$, for some unknown parameter $(\mu, \sigma^2) \in \mathbb{R} \times (0, \infty)$. We want to test the following hypotheses at non-asymptotic level α (for some fixed $\alpha \in (0,1)$):

$$H_0 : \mu \geq 0 \text{ vs. } H_1 : \mu < 0.$$

1. Recall the maximum likelihood estimator $(\hat{\mu}, \hat{\sigma}^2)$ of (μ, σ^2).

2. Let $S = \sqrt{n-1}\frac{\hat{\mu} - \mu}{\sqrt{\hat{\sigma}^2}}$. Prove that S is a Student random variable with $n-1$ degrees of freedom.

3. Propose a test with non-asymptotic level α. Prove your answer.

Non-asymptotic hypothesis test for mean of Gaussians (t-test)

① MLE. Likelihood: $f(\mu, \sigma^2, X_1, \ldots, X_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2}\cdot(X_i - \mu)^2}$

$$\ell(\mu, \sigma^2) = \log f = \sum_{i=1}^{n}\left[-\frac{1}{2}\log 2\pi - \frac{1}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(X_i - \mu)^2\right]$$

$$= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2$$

$$\partial_\mu \ell(\mu, \sigma^2) = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2\cdot(X_i - \mu)(-1) \overset{!}{=} 0$$

$$\Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\partial_{\sigma^2} \ell(\mu, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(X_i - \mu)^2 \overset{!}{=} 0$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\mu})^2$$

① Estimate μ, σ^2

$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\mu})^2$

✓ step 1.

Review: MLE for Gaussian

Going through steps to arrive at T test

Example : adminestering a drug, group of drug test and without drug test

Null hypothesis is drug keeps blood pressure the same or raises it (mu)

$X_1, \ldots, X_n \sim \mathcal{N}(\underbrace{\mu, \sigma^2}_{\text{unknown}})$ iid.

Goal: Test with level α

$H_0: \mu \geq 0 \quad H_1: \mu < 0$

① Estimate μ, σ^2

② Test statistic / pivot

$\psi = \mathbb{1}\{T_n > s\}$

③ Adjust s

$\hat{\mu}$: optimal μ

$\hat{\sigma}^2$ = optimal variance

← multiplied by σ^4

Proving Cochran's Theorem
Orthogonal Matrices

$$② \quad \tilde{T}_n^{(1)} = \sqrt{n}\frac{(\mu - \hat{\mu})}{\sqrt{\sigma^2}} \sim \mathcal{N}(0,1)$$

μ will later be replaced by 0 (don't want test to depend on μ)

σ^2 unknown $\therefore$ replace by $\hat{\sigma}^2$
estimator

what is distribution of $\hat{\sigma}^2$?
(later)

First: Review of orthogonal matrices

$$\underset{\in \mathbb{R}^n}{[v_1} | \cdots | v_n] = V \in \mathbb{R}^{n \times n}$$

$$v_i^T v_j = 0, \; i \neq j, \; v_i^T v_i = \|v\|_2^2 = 1 \Rightarrow \{v_i\}_{i=1,\dots,n} \text{ orthonormal}$$

$$V^T V = \begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_n^T & - \end{bmatrix} \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = I_n = VV^T \quad V \text{ orthogonal}$$

$$x \in \mathbb{R}^n, \; \|Vx\|_2^2 = (Vx)^T(Vx) = x^T \underbrace{V^T V}_{=I_n} x = x^T x = \|x\|_2^2$$

$$W = [\underset{\in \mathbb{R}^n}{w_1} | \cdots | w_k] \in \mathbb{R}^{n \times k}, \; \{w_i\}_{i=1,\dots,k} \text{ orthonormal}$$

$$\|W W^T x\|_2^2 = (WW^T x)^T (WW^T x) = x^T W \underbrace{W^T W}_{I_k} W^T x = \|W^T x\|_2^2$$

now try to understand distribution of $\hat{\sigma}^2$

now try to understand distribution of $\hat{\sigma}^2$

Sample Variance as a Norm

$$② \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\mu})^2$$

$$n\hat{\sigma}^2 = \left\| \underbrace{\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}}_{=X} - \begin{pmatrix} \hat{\mu} \\ \vdots \\ \hat{\mu} \end{pmatrix} \right\|_2^2$$

$$\hat{\mu} = \frac{1}{n} \underset{\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}}{\mathbb{1}^T} X, \quad \begin{pmatrix} \hat{\mu} \\ \vdots \\ \hat{\mu} \end{pmatrix} = \frac{1}{n} \underbrace{\mathbb{1}\mathbb{1}^T}_{=\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}} X$$

$$= \underbrace{\frac{1}{\sqrt{n}}\mathbb{1}}_{=v_1} \underbrace{\left(\frac{1}{\sqrt{n}}\mathbb{1}\right)^T}_{v_1^T} X$$

$$= \|X - v_1 v_1^T X\|_2^2$$

normalise by the mean

$$Y = \frac{X - \mu\mathbb{1}}{\sigma} \sim N(0, I_n)$$

$$= \|X - \mu\mathbb{1} - (\underbrace{v_1 v_1^T X - \mu\mathbb{1}}_{=v_1 v_1^T (X - \mu\mathbb{1})})\|_2^2$$

where $\mu\mathbb{1} = v_1 v_1^T \mu \mathbb{1}$, $v_1 = \frac{1}{\sqrt{n}}\mathbb{1}$

$$\Rightarrow \frac{n\hat{\sigma}^2}{\sigma^2} = \|Y - v_1 v_1^T Y\|_2^2$$

$$V = [v_1 | \cdots | v_n] \in \mathbb{R}^{n \times n} \text{ orthogonal}$$

$$I_n = VV^T = [v_1 | \cdots | v_n]\begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix}$$

$$= \left(\sum_{i=1}^{n} v_i e_i^T\right)\left(\sum_{j=1}^{n} v_j e_j^T\right)^T = \sum_{i,j=1}^{n} v_i \underbrace{e_i^T e_j}_{=\begin{cases} 1, & i=j \\ 0, & i \neq j \end{cases}} v_j^T = \sum_{i=1}^{n} v_i v_i^T$$

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i$$

$$= \left\|\sum_{i=1}^{n} v_i v_i^T Y - v_1 v_1^T Y\right\|_2^2$$

$$= \left\|\sum_{i=2}^{n} v_i v_i^T Y\right\|_2^2$$

$$= \| \sum_{i=2} v_i v_i^T Y \|_2$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

completing proof

Chi Squared Distribution

② $\frac{n\hat{\sigma}^2}{\sigma^2} = \| \sum_{i=2}^{n} v_i v_i^T Y \|_2^2$, $\quad Y = \begin{pmatrix} \frac{X_1 - \mu}{\sigma} \\ \vdots \\ \frac{X_n - \mu}{\sigma} \end{pmatrix} \sim \mathcal{N}(0, I_n)$

$V = [v_1 | \dots | v_n]$ – orthogonal

$= W \in \mathbb{R}^{n \times (n-1)}$

$$= \| W W^T Y \|_2^2 = \| \underbrace{W^T Y}_{\in \mathbb{R}^{n-1}} \|_2^2$$

$$\sum_{i=2}^{n} v_i v_i^T = W W^T$$

$$\Rightarrow \frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-1} \quad \text{chi squared distribution}$$

recall $\hat{\mu} = \frac{1}{\sqrt{n}} v_1^T \cdot X = \mu + \sigma \frac{1}{\sqrt{n}} v_1^T \cdot Y$

$$\mathrm{Cov}(\underbrace{W^T Y}_{\in \mathbb{R}^{n-1}}, \underbrace{v_1^T Y}_{\in \mathbb{R}}) = \mathbb{E}[W^T Y Y^T v_1]$$

$$= W^T \underbrace{\mathbb{E}[Y Y^T]}_{I_n} v_1 = W^T v_1 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{n-1}$$

$$\Rightarrow W^T Y \perp\!\!\!\perp v_1^T Y \Rightarrow \hat{\sigma}^2 \perp\!\!\!\perp \hat{\mu}$$

is independant

$\tilde{T}_n = \hat{\mu} - \mu \quad = \hat{\mu} - \mu \quad$ T distribution with

$$\tilde{T}_n = \frac{\hat{\mu} - \mu}{\sqrt{\hat{\sigma}^2/(n-1)}} = \frac{\hat{\mu} - \mu}{\sqrt{\tilde{\sigma}^2/n}}$$ T distribution with $n-1$ degrees of freedom

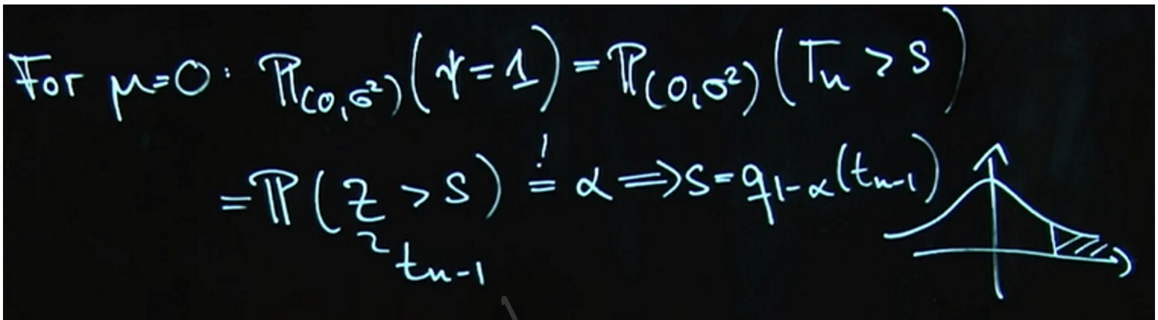
$\tilde{\sigma}^2$ is the unbiased sample variance

$$= \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \hat{\mu})^2$$

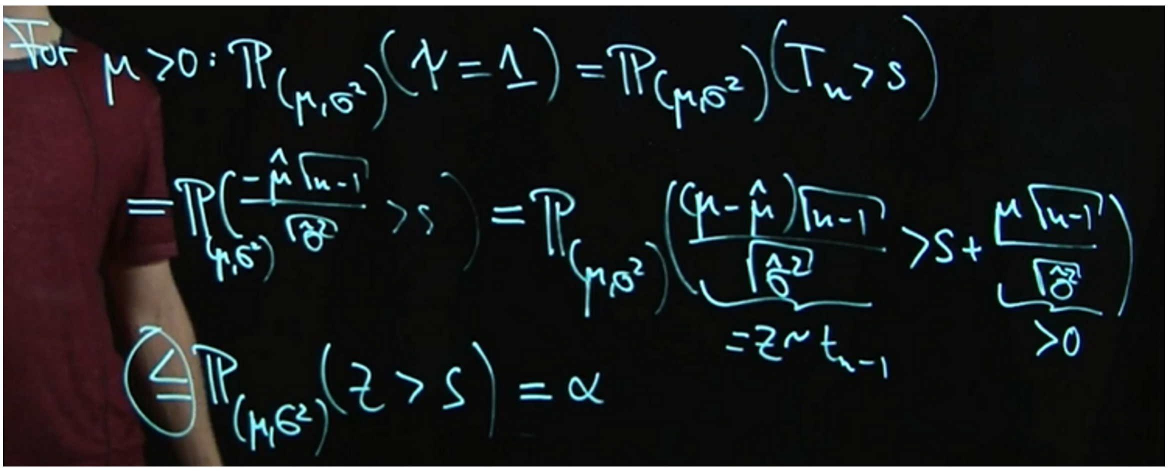
T test conclusion

3.

T distribution looks similar to Gaussian with slightly heavier tails



t with $n-1$ degrees of freedom



type 2 error :