Lecture 4: Sampling Distribution of the MLE of μ and σ^2 for the Normal Distribution

MATH 667-01
Statistical Inference
University of Louisville

August 31, 2017
Last corrected: 9/5/2017

Introduction

- We first state the sampling distribution of the maximum likelihood estimator of the parameters of a normal distribution based on a random sample as stated in Section 5.3 of Casella and Berger (2001)[1].
- We will illustrate the result with an example to assess the performance of the estimator.
- We will review properties of the important normal distribution from Sections 3.3 and 4.6, and we introduce the chi-squared distribution and prove some properties about it as discussed in Section 5.3.
- We will also review some details concerning cumulative distribution functions from Section 1.5, moment generating functions from Sections 2.3 and 4.6, and transformations from Section 4.6.

[1] Casella, G. and Berger, R. (2001). Statistical Inference, second edition. Duxbury Press.

- *Theorem L4.1* (Theorem 5.3.1 on p.218): Let $X_1, \ldots, X_n$ be a random sample from a Normal(μ, σ^2) distribution, and let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$. Then
  - (a) $\bar{X}$ and S^2 are independent random variables.
  - (b) $\bar{X}$ has a Normal$(\mu, \sigma^2/n)$ distribution.
  - (c) $\frac{(n-1)S^2}{\sigma^2}$ has a chi squared distribution with $n-1$ degrees of freedom.
- Let $(\hat{\mu}, \widehat{\sigma^2})$ denote the maximum likelihood estimator of (μ, σ^2). Note that $\bar{X} = \hat{\mu}$ and $\frac{(n-1)S^2}{\sigma^2} = \frac{n\widehat{\sigma^2}}{\sigma^2}$.

Question

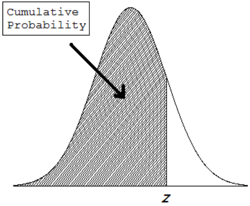
- *Example L4.1*: Suppose that $X_1, \ldots, X_{81}$ be a random sample from a Normal$(1{,}4)$ distribution. What is the probability that both $\hat{\mu}$ and $\widehat{\sigma^2}$ will be within $\frac{1}{2}$ of their true values?

Normal Distribution

- *Definition L4.1* (p.113): The pdf of a random variable X with a *normal* distribution mean μ and variance σ^2 is
$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$
and we write $X \sim \text{Normal}(\mu,\sigma^2)$.
- *Definition L4.2* (p.113): A normal random variable is called *standard normal* if $\mu = 0$ and $\sigma^2 = 1$.
- If $X \sim \text{Normal}(\mu,\sigma^2)$, then $Z = \dfrac{X-\mu}{\sigma} \sim \text{Normal}(0,1)$.
- *Theorem L4.2* (Cor 4.6.10 on p.184): Let $X_1,\ldots,X_n$ be mutually independent random variables with $X_i \sim \text{Normal}(\mu_i,\sigma_i^2)$. Let $a_1,\ldots,a_n$ and $b_1,\ldots,b_n$ be fixed constants. Then
$$\sum_{i=1}^{n}(a_iX_i + b_i) \sim \text{Normal}\left(\sum_{i=1}^{n}(a_i\mu_i + b_i), \sum_{i=1}^{n} a_i^2\sigma_i^2\right).$$

Normal Distribution

cdf of standard normal distribution: $F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$



Standard Normal Cumulative Probabilities (continued)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
⋮										
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952

Chi Squared Distribution

- *Definition L4.3* (p.101): The pdf of a random variable Y with a *chi squared* distribution with p degrees of freedom is

$$f(y|p) = \frac{1}{\Gamma(p/2)2^{p/2}} y^{(p/2)-1} e^{-y/2} I_{(0,\infty)}(y)$$

  and we write $Y \sim \chi^2_p$.

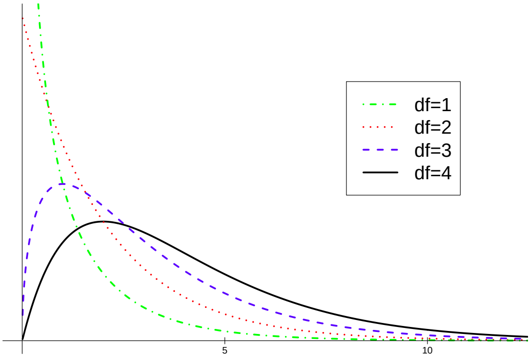
- The chi squared distribution is a special case of the gamma distribution (p.99) with $\alpha = p/2$ and $\beta = 2$. The pdf of a gamma(α,β) distribution is

$$f(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta} I_{(0,\infty)}(x)$$

  where $\alpha > 0$ is a shape parameter and $\beta > 0$ is a scale parameter.

- If $Y \sim \chi^2_p$, then E$[Y] = \alpha\beta = (\frac{p}{2})2 = p$ and Var$[Y] = \alpha\beta^2 = (\frac{p}{2})2^2 = 2p$.

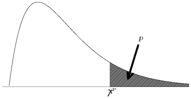
pdf of chi squared distribution



Chi Squared Distribution

Table of tail probabilities for χ^2 distribution

$$\int_{\chi^{2*}}^{\infty} \frac{1}{\Gamma(df/2)2^{df/2}} y^{(df/2)-1} e^{-y/2}\ dy = p$$



χ^2 distribution critical values

	Upper-tail probability p							
df	.25	.20	.15	.10	.05	.025	.02	.01
1	1.323	1.642	2.072	2.706	3.841	5.024	5.412	6.635
2	2.773	3.219	3.794	4.605	5.991	7.378	7.824	9.210
3	4.108	4.642	5.317	6.251	7.815	9.348	9.837	11.345
				⋮				
60	66.981	68.972	71.341	74.397	79.082	83.298	84.580	88.379
80	88.130	90.405	93.106	96.578	101.879	106.629	108.069	112.329
100	109.141	111.667	114.659	118.498	124.342	129.561	131.142	135.807

Question

- *Example L4.1*: Suppose that $X_1, \ldots, X_{81}$ be a random sample from a Normal$(1,4)$ distribution. What is the probability that both $\hat{\mu}$ and $\widehat{\sigma^2}$ will be within $\frac{1}{2}$ of their true values?
- *Answer to Example L4.1*:

$$
\begin{aligned}
P\left(|\bar{X}-1| < \frac{1}{2}\right) &= P\left(-\frac{1}{2} < \bar{X}-1 < \frac{1}{2}\right) \\
&= P\left(-\frac{9}{4} < \frac{\bar{X}-1}{2/9} < \frac{9}{4}\right) \\
&= P(-2.25 < Z < 2.25) \\
&= P(Z < 2.25) - P(Z \leq -2.25) \\
&\approx .9878 - (1 - .9878) = .9756
\end{aligned}
$$

This can be computed with the R command

```
pnorm(2.25)-pnorm(-2.25)
```

- *Answer to Example L4.1 continued*:

$$
\begin{aligned}
P\left(\left|\widehat{\sigma^2}-4\right|<\frac{1}{2}\right) &= P\left(3.5<\widehat{\sigma^2}<4.5\right) \\
&= P\left(70.875<\frac{81\widehat{\sigma^2}}{4}<91.125\right) \\
&= \int_{70.875}^{91.125}\frac{1}{\Gamma(40)2^{40}}y^{39}e^{-y/2}dy \\
&= P\left(\frac{81\widehat{\sigma^2}}{4}<91.125\right)-P\left(\frac{81\widehat{\sigma^2}}{4}\le 70.875\right) \\
&\approx .8143-(1-.2427)=.5716
\end{aligned}
$$

This can be computed with the R command

```
pchisq(91.125,df=80)-pchisq(70.875,df=80)
```

- *Answer to Example L4.1 continued*:

$$
\begin{aligned}
&P\left(|\bar{X}-1| < \frac{1}{2} \text{ and } \left|\widehat{\sigma^2}-4\right| < \frac{1}{2}\right)\\
&= P\left(|\bar{X}-1| < \frac{1}{2}\right) P\left(\left|\widehat{\sigma^2}-4\right| < \frac{1}{2}\right)\\
&= \int_{-2.25}^{2.25} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \int_{70.875}^{91.125} \frac{1}{\Gamma(40)2^{40}} y^{39} e^{-y/2} dy\\
&\approx .5576
\end{aligned}
$$

- *Theorem L4.3* (Lem 5.3.2 on p.219): We use the notation χ^2_p to denote a chi squared random variable with p degrees of freedom.
  (a) If Z is a Normal$(0,1)$ random variable, then $Z^2 \sim \chi^2_1$; that is, the square of a standard normal random variable is a chi squared random variable.
  (b) If $X_1, \ldots, X_n$ are independent, and $X_i \sim \chi^2_{p_i}$, then $X_1 + \cdots + X_n \sim \chi^2_{p_1+\cdots+p_n}$; that is, independent chi squared random variables add to a chi squared random variable, and the degrees of freedom also add.

Review of Cumulative Distribution Functions

- *Definition L4.4* (Def 1.5.1 on p.29): The *cumulative distribution function* (cdf) of X is $F_X(x) = P_X(X \leq x)$ for all x.
- *Theorem L4.4* (Thm 1.5.3 on p.31): The function $F(x)$ is a cdf if and only if the following conditions hold:
  (a) $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$
  (b) $F(x)$ is nondecreasing in x
  (c) $F(x)$ is right continuous; that is, for every x_0, $\lim_{x \downarrow x_0} F(x) = F(x_0)$.
- *Definition L4.5* (Def 1.5.8 on p.33): Random variables X and Y are *identically distributed* if for every set A, $P(X \in A) = P(Y \in A)$.
- *Theorem L4.5* (Thm 1.5.10 on p.34): X and Y are identically distributed if and only if $F_X(x) = F_Y(x)$ for every x.
- (Eq 1.6.2 on p.35): For a continuous random variable X with cdf F and pdf f, $\dfrac{dF(x)}{dx} = f(x)$.

Chi Squared Distribution

- *Proof of Theorem L4.3(a)*: For $y \geq 0$, the cdf of $Y = Z^2$ is

$$\begin{aligned}
F_Y(y) &= P\left(Y \leq y\right) \\
&= P\left(Z^2 \leq y\right) \\
&= P\left(-\sqrt{y} \leq Z \leq \sqrt{y}\right) \\
&= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz.
\end{aligned}$$

Then, for $y \geq 0$, the pdf of Y is

$$\begin{aligned}
f_Y(y) = \frac{dF_Y(y)}{dy} &= \frac{d}{dy}\left[\int_0^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz - \int_0^{-\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz\right] \\
&= \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{y})^2/2} \frac{d}{dy}\left[\sqrt{y}\right] - \frac{1}{\sqrt{2\pi}} e^{-(-\sqrt{y})^2/2} \frac{d}{dy}\left[-\sqrt{y}\right] \\
&= \frac{1}{\sqrt{2\pi}} e^{-y/2} \frac{1}{2\sqrt{y}} - \frac{1}{\sqrt{2\pi}} e^{-y/2}\left(-\frac{1}{2\sqrt{y}}\right) = \frac{1}{2^{1/2}\Gamma(\frac{1}{2})} e^{-y/2} y^{-1/2}.
\end{aligned}$$

Review of Moment Generating Functions

- *Definition L4.6* (Def 2.3.6 on p.62): Let X be a random variable with cdf F_X. The *moment generating function (mgf) of* X (or F_X) is $M_X(t) = \mathsf{E}[e^{tX}]$, provided that the expectation exists for t in some neighborhood of 0 (that is, there is an $h > 0$ such that $\mathsf{E}[e^{tX}]$ exists for all t in $-h < t < h$).
- *Theorem L4.6* (Thm 2.3.11 on p.65): Let X and Y be random variables with moment generating functions M_X and M_Y, respectively. If $M_X(t) = M_Y(t)$ for all t in some neighborhood of 0, then $F_X(u) = F_Y(u)$ for all u.
- *Theorem L4.7* (Thm 4.6.7 on p.183): Let $X_1, \ldots, X_n$ be mutually independent random variables with mgfs $M_{X_1}(t), \ldots, M_{X_n}(t)$. Let $Z = \sum_{i=1}^{n} X_i$. Then the mgf of Z is $M_Z(t) = \prod_{i=1}^{n} M_{X_i}(t)$.

- *Proof of Theorem L4.3(b)*:
  The moment generating function of a Gamma(α, β) random variable is $M(t) = \left(\frac{1}{1-\beta t}\right)^{\alpha}$ for $t < \frac{1}{\beta}$.
  Since the chi squared distribution with p_i degrees of freedom is the Gamma($\frac{p_i}{2}$,2) distribution, let $\alpha_i = \frac{p_i}{2}$ and $\beta = 2$.
  If $X_i \sim$ Gamma(α_i, β), then

$$M_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n M_{X_i}(t) = \prod_{i=1}^n \left(\frac{1}{1-\beta t}\right)^{\alpha_i} = \left(\frac{1}{1-\beta t}\right)^{\sum_{i=1}^n \alpha_i}.$$

  This implies that $\sum_{i=1}^n X_i \sim$ Gamma$(\sum_{i=1}^n \alpha_i, \beta)$.
  Since $\sum_{i=1}^n \alpha_i = \frac{1}{2}\sum_{i=1}^n p_i$ and $\beta = 2$, it follows that $\sum_{i=1}^n X_i \sim \chi^2_{\sum_{i=1}^n p_i}$.

Review of Multivariate Transformations

- Multivariate continuous case (p.185): Suppose $\boldsymbol{X} = (X_1, \ldots, X_n)$ has pdf $f(x_1, \ldots, x_n)$ and $U_i = u_i(x_1, \ldots, x_n), i = 1, \ldots, n$ where u_i has an inverse such that there is an h_i where $x_i = h_i(u_1, \ldots, u_n)$. Then the joint pdf of $\boldsymbol{U} = (U_1, \ldots, U_n)$ is

$$g(u_1, \ldots, u_n) = f(h_1(u_1, \ldots, u_n), \ldots, h_n(u_1, \ldots, u_n)) \, |J|$$

where

$$J = \left| \frac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, u_n)} \right| = \begin{vmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \cdots & \frac{\partial x_2}{\partial u_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial u_1} & \frac{\partial x_n}{\partial u_2} & \cdots & \frac{\partial x_n}{\partial u_n} \end{vmatrix}.$$

Main Result

- *Theorem L4.1* (Theorem 5.3.1 on p.218): Let $X_1, \ldots, X_n$ be a random sample from a Normal(μ, σ^2) distribution, and let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$. Then
  (a) $\bar{X}$ and S^2 are independent random variables.
  (b) $\bar{X}$ has a Normal$(\mu, \sigma^2/n)$ distribution.
  (c) $\frac{(n-1)S^2}{\sigma^2}$ has a chi squared distribution with $n-1$ degrees of freedom.
- Let $(\hat{\mu}, \widehat{\sigma^2})$ denote the maximum likelihood estimator of (μ, σ^2). Note that $\bar{X} = \hat{\mu}$ and $\frac{(n-1)S^2}{\sigma^2} = \frac{n\widehat{\sigma^2}}{\sigma^2}$.

Proof of Theorem L4.1(a):

- Let $Z_i = \dfrac{X_i - \mu}{\sigma}$. Then $Z_1, \ldots, Z_n$ are independent Normal$(0,1)$ random variables (see Theorem L3.5 and slide 4.5).
- $\bar{Z} = \dfrac{\sum_{i=1}^n Z_i}{n} = \dfrac{\bar{X} - \mu}{\sigma}$ and

$$S_Z^2 = \frac{1}{n-1}\sum_{i=1}^n (Z_i - \bar{Z})^2 = \frac{1}{(n-1)\sigma^2}\sum_{i=1}^n (X_i - \bar{X})^2 \overset{3.22}{=} \frac{S^2}{\sigma^2}$$

- Now we want to show that $\bar{Z}$ and S_Z^2 are independent. We will do this by showing that $U_1 = \bar{Z}$ and $(U_2, \ldots, U_n) = (Z_2 - \bar{Z}, \ldots, Z_n - \bar{Z})$ are independent using Theorem L3.5.

Proof of Theorem L4.1(a) continued:

- Note that

$$
\begin{aligned}
S_Z^2 &= \frac{1}{n-1}\left[(Z_1-\bar{Z})^2+\sum_{i=2}^{n}(Z_i-\bar{Z})^2\right] \\
&= \frac{1}{n-1}\left[(-\sum_{i=2}^{n}(Z_i-\bar{Z}))^2+\sum_{i=2}^{n}(Z_i-\bar{Z})^2\right] \\
&= \frac{1}{n-1}\left[(\sum_{i=2}^{n}U_i)^2+\sum_{i=2}^{n}U_i^2\right]
\end{aligned}
$$

can be expressed as a function of $U_2,\ldots,U_n$.

Proof of Theorem L4.1(a) continued:

- The transformation from $z_1, \ldots, z_n$ to $u_1, \ldots, u_n$ can be inverted as follows:
$$\left.\begin{array}{l} u_1 = \bar{z} \\ u_i = z_i - \bar{z}, i = 2, \ldots, n \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} z_1 = u_1 - \sum_{i=2}^n u_i \\ z_i = u_i + u_1, i = 2, \ldots, n \end{array}\right. .$$
- Since the joint pdf of $Z_1, \ldots, Z_n$ is $f_{\boldsymbol{Z}}(z_1, \ldots, z_n) = \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^n z_i^2\right\}$, the joint density of U_1 and $(U_2, \ldots, U_n)$ is

$$\begin{aligned} f_{\boldsymbol{U}}(u_1, u_2, \ldots, u_n) &= f_{\boldsymbol{Z}}(u_1 - \sum_{i=2}^n u_i, u_2 + u_1, \ldots, u_n + u_1) \left|\det\left(\frac{\partial(z_1, \ldots, z_n)}{\partial(u_1, \ldots, u_n)}\right)\right| \\ &= \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}(u_1 - \sum_{i=2}^n u_i)^2\right\} \exp\left\{-\frac{1}{2}\sum_{i=2}^n (u_i + u_1)^2\right\} \cdot n \\ &= \frac{n}{(2\pi)^{n/2}} e^{-nu_1^2/2} \exp\left\{-\frac{1}{2}\left[\sum_{i=2}^n u_i^2 + \left(\sum_{i=2}^n u_i\right)^2\right]\right\} \end{aligned}$$

Proof of Theorem L4.1(a) continued:

since
$$\left|\frac{\partial(z_1,\ldots,z_n)}{\partial(u_1,\ldots,u_n)}\right| = \begin{vmatrix} 1 & -1 & -1 & -1 & \cdots & -1 \\ 1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ & & & \vdots & & \end{vmatrix}$$
$$\overset{R_i+R_1\to R_1,i=2,\ldots,n}{=} \begin{vmatrix} n & 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ & & & \vdots & & \end{vmatrix} = n$$

- Since $f_{\boldsymbol{U}}(u_1, u_2, \ldots, u_n) = g_1(u_1) \cdot g_2(u_2, \ldots, u_n)$, U_1 and $(U_2, \ldots, U_n)$ are independent by Theorem L3.4.

Proof of Theorem L4.1(a) continued:

- Thus, since $\bar{Z} = U_1$ is a function of U_1 and S_Z^2 is a function of $U_2, \ldots, U_n$, $\bar{Z}$ and S_Z^2 are independent by Theorem L3.5.
- Furthermore, since $\bar{X}$ is a function of $\bar{Z}$ and S^2 is a function of S_Z^2, $\bar{X}$ and S^2 are also independent by Theorem L3.5.

Proof of Theorem L4.1(b):

- The mgf of a Normal(μ, σ^2) random variable is $M_{X_i}(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$. So, the mgf of $\bar{X}$ is
$$M_{\bar{X}}(t) = \mathsf{E}[e^{t\bar{X}}] = \mathsf{E}[e^{\frac{t}{n}\sum_{i=1}^n X_i}] = \mathsf{E}\left[\prod_{i=1}^n e^{\frac{t}{n}X_i}\right]$$
$$\overset{3.6}{=} \prod_{i=1}^n \mathsf{E}\left[e^{\frac{t}{n}X_i}\right] = \prod_{i=1}^n e^{\mu(t/n)+\frac{1}{2}\sigma^2(t/n)^2}$$
$$= \left(e^{\mu(t/n)+\frac{1}{2}\sigma^2(t/n)^2}\right)^n = e^{\mu t + \frac{1}{2}\frac{\sigma^2}{n}t^2}.$$
- So, Theorem L4.6 and Theorem L4.5 imply that $\bar{X} \sim$ Normal$(\mu, \frac{\sigma^2}{n})$.

Proof of Theorem L4.1(c):

- Let $\bar{Z}_k = \frac{\sum_{i=1}^{k} Z_i}{k}$ and $S_k^2 = \frac{1}{k-1}\sum_{i=1}^{k}(Z_i - \bar{Z}_k)^2$.
- $\bar{Z}_{n+1} = \dfrac{\sum_{i=1}^{n+1} Z_i}{n+1} = \dfrac{Z_{n+1} + \sum_{i=1}^{n} Z_i}{n+1} = \dfrac{Z_{n+1} + n\bar{Z}_n}{n+1}$
- For $n \geq 2$, we now show that
  $nS_{n+1}^2 = (n-1)S_n^2 + \frac{n}{n-1}(Z_{n+1} - \bar{Z}_n)^2$.

$$
\begin{aligned}
nS_{n+1}^2 &\overset{3.18}{=} n\left[\frac{1}{(n+1-1)}\left\{\sum_{i=1}^{n+1} Z_i^2 - (n+1)\bar{Z}_{n+1}^2\right\}\right] \\
&= \sum_{i=1}^{n+1} Z_i^2 - (n+1)\bar{Z}_{n+1}^2 \\
&= \sum_{i=1}^{n} Z_i^2 - n\bar{Z}_n^2 + Z_{n+1}^2 - (n+1)\bar{Z}_{n+1}^2 + n\bar{Z}_n^2 \\
&\overset{3.18}{=} (n-1)S_n^2 + Z_{n+1}^2 - (n+1)\bar{Z}_{n+1}^2 + n\bar{Z}_n^2
\end{aligned}
$$

Proof of Theorem L4.1(c) continued:

$$\overset{3.26}{=} (n-1)S_n^2 + Z_{n+1}^2 - (n+1)\frac{(Z_{n+1} + n\bar{Z}_n)^2}{(n+1)^2} + n\bar{Z}_n^2$$

$$= (n-1)S_n^2 + Z_{n+1}^2 - (n+1)\frac{Z_{n+1}^2 + 2nZ_{n+1}\bar{Z}_n + n^2\bar{Z}_n^2}{(n+1)^2} + \frac{n(n+1)\bar{Z}_n^2}{n+1}$$

$$= (n-1)S_n^2 + \frac{nZ_{n+1}^2 - 2nZ_{n+1}\bar{Z}_n + n\bar{Z}_n^2}{n+1}$$

$$= (n-1)S_n^2 + \frac{n}{n-1}(Z_{n+1}^2 - 2Z_{n+1}\bar{Z}_n + \bar{Z}_n^2)$$

$$= (n-1)S_n^2 + \frac{n}{n-1}(Z_{n+1} - \bar{Z}_n)^2$$

Proof of Theorem L4.1(c) continued:

- Next, we will show that $(n-1)S_n^2 \sim \chi^2_{n-1}$ by induction.
- *Basis Step (Show for $n=2$)*:

$$\begin{aligned} S_2^2 &= (Z_1 - \frac{1}{2}Z_1 - \frac{1}{2}Z_2)^2 + (Z_2 - \frac{1}{2}Z_1 - \frac{1}{2}Z_2)^2 \\ &= \left(\frac{1}{2}Z_1 - \frac{1}{2}Z_2\right)^2 + \left(\frac{1}{2}Z_2 - \frac{1}{2}Z_1\right)^2 \\ &= 2\left(\frac{1}{2}Z_1 - \frac{1}{2}Z_2\right)^2 \\ &= \left(\frac{1}{\sqrt{2}}Z_1 - \frac{1}{\sqrt{2}}Z_2\right)^2 \end{aligned}$$

Since $\frac{1}{\sqrt{2}}Z_1 - \frac{1}{\sqrt{2}}Z_2 \sim \text{Normal}(0,1)$ (see Theorem L4.2), $S_2^2 = \frac{1}{2}(Z_2 - Z_1)^2 = \left(\frac{1}{\sqrt{2}}Z_1 - \frac{1}{\sqrt{2}}Z_2\right)^2 \sim \chi^2_1$ by Theorem L4.3(a).

Proof of Theorem L4.1(c) continued:

- *Inductive Step*:
  Now assume $(k-1)S_k^2 \sim \chi^2_{k-1}$ and show that $kS_{k+1}^2 \sim \chi^2_k$.
  Z_{k+1} is independent of $(Z_1, \ldots, Z_k)$ by Theorem L3.5.
  $\bar{Z}_k$ and Z_k^2 are independent by Theorem L4.1(a).
  So, Z_{k+1},$\bar{Z}_k$, and Z_k^2 are mutually independent.
  By Theorem L3.5, $\frac{k}{k+1}(Z_{k+1} - \bar{Z}_k)^2$ is independent of $(k-1)S_k^2$.
  Since $Z_{k+1} - \bar{Z}_k \sim \text{Normal}(0, 1 + \frac{1}{k})$,
  $\sqrt{\frac{k}{k+1}}(Z_{k+1} - \bar{Z}_k) \sim \text{Normal}(0, 1)$ so that
  $\frac{k}{k+1}(Z_{k+1} - \bar{Z}_k)^2 \sim \chi^2_1$ by Theorem L4.3(a).
  So, $kS_{k+1}^2 = (k-1)S_k^2 + \frac{k}{k+1}(Z_{k+1} - \bar{Z}_k)^2 \sim \chi^2_k$ by Theorem L4.3(b).
- Thus, it has been shown that $(n-1)S_Z^2 \sim \chi^2_{n-1}$. Since $S_Z^2 \overset{5.22}{=} \frac{S^2}{\sigma^2}$, we have the desired result: $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$.