MTH375: Mathematical Statistics - Homework #1

Cason Konzer
github

Key Concepts: Multivariate normal($\boldsymbol{\mu}, \Sigma$), χ^2 with n df, t with n df, F with m, n df

1. Suppose that Z_1, Z_2 are independent standard normal random variables.

Let $Y_1 = Z_1 - 2Z_2$, $Y_2 = Z_1 - Z_2$.

(a) Find the joint pdf $f_{Y_1,Y_2}(y_1, y_2)$.

You may do this either of two ways. Either (i) use the change of variables theorem from MTH 372, OR (ii) evaluate the matrices Σ and Σ^{-1} from class, then multiply the necessary matrices and vectors to obtain a formula for $f_{Y_1,Y_2}(y_1, y_2)$. In either case, obtain a formula containing no matrices and no vectors.

Solution:

We have a system of equations, $\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = A\boldsymbol{Z} + \boldsymbol{\mu} = \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

We can solve for $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}} = \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} = \boldsymbol{\Sigma}$

We can now evaluate $\boldsymbol{\Sigma}^{-1}$

$$\begin{bmatrix} 5 & 3 & 1 & 0 \\ 3 & 2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 6 & 2 & 0 \\ 3 & 2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 & -3 \\ 3 & 2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 & -3 \\ 0 & 2 & -6 & 10 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 & -3 \\ 0 & 1 & -3 & 5 \end{bmatrix}$$

Our solution yields $\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}$

The last piece we need to evaluate before solving for the joint pdf is $||\boldsymbol{\Sigma}||$

$$||\boldsymbol{\Sigma}|| = \left|\left|\begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}\right|\right| = |10 - 9| = |1| = 1$$

We can now solve for $f_{Y_1,Y_2}(y_1, y_2) = \dfrac{e^{-\frac{(\boldsymbol{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}-\boldsymbol{\mu})}{2}}}{(2\pi)^{\frac{n}{2}} \sqrt{||\boldsymbol{\Sigma}||}} = \dfrac{e^{-\frac{(\boldsymbol{y})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{y})}{2}}}{(2\pi)^{\frac{2}{2}} \sqrt{1}} = \dfrac{e^{-\frac{(\boldsymbol{y})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{y})}{2}}}{2\pi}$

$$\boldsymbol{y}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2y_1 - 3y_2 & -3y_1 + 5y_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= 2y_1^2 - 3y_2y_1 - 3y_1y_2 + 5y_2^2 = 2y_1^2 - 6y_1y_2 + 5y_2^2$$

By Substitution, $f_{Y_1,Y_2}(y_1, y_2) = \dfrac{1}{2\pi} \exp\left(\dfrac{-2y_1^2 + 6y_1y_2 - 5y_2^2}{2} \right)$

(b) Find the marginal pdf $f_{Y_2}(y_2)$. Don't use integration – you can derive the needed pdf doing only the simplest arithmetic using some facts from MTH 372.

Solution:

We have one equation, $\boldsymbol{Y} = Y_2 = A\boldsymbol{Z} + \boldsymbol{\mu} = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = Z_1 - Z_2$

We can solve for $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A^T} = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2 = \boldsymbol{\Sigma}$

We can now evaluate $\boldsymbol{\Sigma}^{-1} = \dfrac{1}{2}$

The last piece we need to evaluate before solving for the joint pdf is $||\boldsymbol{\Sigma}|| = 2$

We can now solve for $f_{Y_2}(y_2) = \dfrac{e^{-\frac{(\boldsymbol{y}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})}{2}}}{(2\pi)^{\frac{n}{2}}\sqrt{||\boldsymbol{\Sigma}||}} = \dfrac{e^{-\frac{(\boldsymbol{y})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y})}{2}}}{(2\pi)^{\frac{1}{2}}\sqrt{2}} = \dfrac{e^{-\frac{(\boldsymbol{y})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y})}{2}}}{\sqrt{2}\sqrt{2\pi}}$

$\boldsymbol{y}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} = y_2\dfrac{1}{2}y_2 = \dfrac{y_2^2}{2}$

By Substitution, $f_{Y_2}(y_2) = \dfrac{1}{\sqrt{2}\sqrt{2\pi}}\exp\left(\dfrac{-y_2^2}{4}\right)$

(c) Find the conditional pdf of Y_1 given $Y_2 = 1$, that is, $f_{Y_1|Y_2}(y_1|1)$. Use (a) and (b), and do the necessary division. By completing the square, identify by name – *including parameters* – the required conditional pdf.

Solution:

$$f_{Y_1|Y_2}(y_1|y_2) = \frac{f_{Y_1,Y_2}(y_1,y_2)}{f_{Y_2}(y_2)} = \frac{\frac{1}{2\pi}\exp\left(\frac{-2y_1^2+6y_1y_2-5y_2^2}{2}\right)}{\frac{1}{\sqrt{2}\sqrt{2\pi}}\exp\left(\frac{-y_2^2}{4}\right)} = \frac{\sqrt{2}\exp\left(\frac{-2y_1^2+6y_1y_2-5y_2^2}{2}\right)}{\sqrt{2\pi}\exp\left(\frac{-y_2^2}{4}\right)}$$

$$= \frac{\sqrt{2}\exp\left(\frac{-4y_1^2+12y_1y_2-10y_2^2}{4} - \frac{-y_2^2}{4}\right)}{\sqrt{2\pi}} = \frac{\sqrt{2}\exp\left(\frac{-4y_1^2+12y_1y_2-9y_2^2}{4}\right)}{\sqrt{2\pi}}$$

$$f_{Y_1|Y_2}(y_1|y_2=1) = \frac{f_{Y_1,Y_2}(y_1,y_2=1)}{f_{Y_2}(y_2=1)} = \frac{\sqrt{2}\exp\left(\frac{-4y_1^2+12y_1(1)-9(1)^2}{4}\right)}{\sqrt{2\pi}} = \frac{\sqrt{2}\exp\left(\frac{-4y_1^2+12y_1-9}{4}\right)}{\sqrt{2\pi}}$$

$$= \frac{\sqrt{2}\exp\left(-(\frac{1}{2})(\frac{4y_1^2-12y_1+9}{2})\right)}{\sqrt{2\pi}} = \frac{\sqrt{2}\exp\left(-(\frac{1}{2})(2y_1^2-6y_1+4.5)\right)}{\sqrt{2\pi}} = \frac{\sqrt{2}\exp\left(-(y_1^2-3y_1+2.25)\right)}{\sqrt{2\pi}}$$

$$= \frac{\sqrt{2}\exp\left(-(y_1-1.5)^2\right)}{\sqrt{2\pi}}$$

It is now worthwhile to considerd the pdf of Random Variable $X \sim N(\mu, \sigma^2)$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$

Under close inspection, we can see that our above distribution for $f_{Y_1|Y_2}(y_1|y_2=1)$ is normal, taking $\sigma = \frac{1}{\sqrt{2}}$, $\sigma^2 = \frac{1}{2}$, and $\mu = 1.5$

---

2. Suppose $Z_1, Z_2, \ldots, Z_6$ are independent standard normal random variables.

(a) Find a number a such that $P(-a < 3Z_1 + 2Z_2 - 4Z_3 < a) = .99$.

Solution:

Let $A = 3Z_1 + 2Z_2 - 4Z_3$

$A \sim N(0, (3^2 + 2^2 + 4^2)) \sim N(0, (9 + 4 + 16)) \sim N(0, 29)$

We now wish to have probabilities for $-a$ corresponding to 0.005 and a corresponding to 0.995.

```
qnorm(p = c(0.005, 0.995), mean = 0, sd = 29)
```

No. Variance = 29, sd = sqrt(29).

```
## [1] -74.69905  74.69905
```

We can see here that $a = 74.69905$ is the solution. Much too big. 1.5/2

(b) Find numbers a, b such that $P(a < Z_1^2 + Z_2^2 + \cdots + Z_6^2 < b) = .99$.

Solution:

Let $B = Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2 + Z_5^2 + Z_6^2$
$B \sim \chi_6^2$

Let us make the generalizing assumption $a = 0$.

We now wish to have probabilities for a corresponding to 0.0 and b corresponding to 0.99.

```
qchisq(p = c(0.0, 0.99), df = 6)
```

```
## [1]  0.00000 16.81189
```

We can see here that $a = 0.0$ and $b = 16.81189$ is one solution.

(c) Find a numbers a such that $P\left(-a < \dfrac{Z_1}{\sqrt{Z_2^2 + Z_3^2 + \cdots + Z_6^2}} < a\right) = .99$.

Solution:

Let $C = \dfrac{Z_1}{\sqrt{Z_2^2 + Z_3^2 + \cdots + Z_6^2}}$

We will leverage the following two points to solve this problem:

- $C\sqrt{6} = \dfrac{Z_1\sqrt{6}}{\sqrt{Z_2^2 + Z_3^2 + \cdots + Z_6^2}} \sim T_6$
- $P\left(-a < \dfrac{Z_1}{\sqrt{Z_2^2 + Z_3^2 + \cdots + Z_6^2}} < a\right) = P\left(-a\sqrt{6} < \dfrac{Z_1\sqrt{6}}{\sqrt{Z_2^2 + Z_3^2 + \cdots + Z_6^2}} < a\sqrt{6}\right)$

We now wish to have probabilities for $-a\sqrt{6}$ corresponding to 0.005 and $a\sqrt{6}$ corresponding to 0.995.

In particular, we will solve $P\left(-a\sqrt{6} < T_6 < a\sqrt{6}\right) = 0.99$.

```
a_sqrt_6 <- qt(p = c(0.005, 0.995), df = 6)
a_sqrt_6
```

```
## [1] -3.707428  3.707428
```

```
a_sqrt_6 / sqrt(6)
```

```
## [1] -1.513551  1.513551
```

We can see here that $a = 1.513551$ is the solution.

5 df, not 6. (We can't use Z1 in the denominator, because it must be independent of the numerator.)

1.5/2

(d) Find numbers a, b such that: $P\left(a < \dfrac{Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2}{Z_5^2 + Z_6^2} < b\right) = .99$.

Solution:

Let $D = \dfrac{Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2}{Z_5^2 + Z_6^2}$

We will leverage the following two points to solve this problem:

- $\dfrac{D}{2} = \dfrac{1}{2} \cdot \dfrac{Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2}{Z_5^2 + Z_6^2} = \dfrac{2\chi_4}{4\chi_2} \sim F_{4,2}$
- $P\left(a < \dfrac{Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2}{Z_5^2 + Z_6^2} < b\right) = P\left(\dfrac{a}{2} < \dfrac{1}{2} \cdot \dfrac{Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2}{Z_5^2 + Z_6^2} < \dfrac{b}{2}\right)$

Let us make the generalizing assumption $a = 0$.

We now wish to have probabilities for a corresponding to 0.0 and b corresponding to 0.99.

In particular, we will solve $P\left(0 < F_{4,2} < \dfrac{b}{2}\right) = 0.99$.

```
ab_halved <- qf(p = c(0.0, 0.99), df1 = 4, df2 = 2)
ab_halved
```

```
## [1]  0.00000 99.24937
```

```
ab_halved * 2
```

```
## [1]   0.0000 198.4987
```

We can see here that $a = 0.0$ and $b = 198.4987$ is one solution.

---

3. Let $X_1, X_2, \ldots, X_8$ be independent normal(μ, σ^2) random variables,

$\overline{X} = \dfrac{X_1 + X_2 + \ldots + X_8}{8}$, and $S^2 = \dfrac{1}{7}\sum_{i=1}^{8}(X_i - \overline{X})^2$.

Find a number a such that $P\left(-a < \dfrac{\overline{X} - \mu}{S} < a\right) = .99$.

Solution:

We will leverage the following from class in this solution...

- $\overline{X} \sim N(\mu, \dfrac{\sigma^2}{n})$
- $\dfrac{(n-1)S^2}{\sigma^2} = \dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2 \sim \chi^2_{n-1}$
- $\overline{X}$ and S are independent

Letting $Y = \dfrac{\overline{X} - \mu}{S}$

Under close inspection, for arbitrary n, we can see the following...

$$Y \cdot \sqrt{n} = \frac{\overline{X} - \mu}{S} \cdot \sqrt{n} \cdot \frac{\sigma}{\sigma} \cdot \frac{\sqrt{n-1}}{\sqrt{n-1}} = \frac{(\overline{X} - \mu)\sqrt{n}}{\sigma} \cdot \frac{\sqrt{\sigma^2}}{\sqrt{S^2}\sqrt{n-1}} \cdot \frac{\sqrt{n-1}}{1} \sim Z \cdot \frac{\sqrt{n-1}}{\sqrt{\chi^2_{n-1}}} \sim T_{n-1}$$

We will thus leverage the following two points to solve this problem:

- $Y\sqrt{8} = \dfrac{(\overline{X} - \mu)\sqrt{8}}{\sigma} \cdot \dfrac{\sqrt{\sigma^2}}{\sqrt{S^2}\sqrt{7}} \cdot \dfrac{\sqrt{7}}{1} \sim Z \cdot \dfrac{\sqrt{7}}{\sqrt{\chi^2_7}} \sim T_7$
- $P\left(-a < \dfrac{\overline{X} - \mu}{S} < a\right) = P\left(-a\sqrt{8} < \dfrac{(\overline{X} - \mu)\sqrt{8}}{\sigma} \cdot \dfrac{\sqrt{\sigma^2}}{\sqrt{S^2}\sqrt{7}} \cdot \dfrac{\sqrt{7}}{1} < a\sqrt{8}\right)$

We now wish to have probabilities for $-a\sqrt{8}$ corresponding to 0.005 and $a\sqrt{8}$ corresponding to 0.995.

In particular, we will solve $P\left(-a\sqrt{8} < T_7 < a\sqrt{8}\right) = 0.99$.

```r
a_sqrt_8 <- qt(p = c(0.005, 0.995), df = 7)
a_sqrt_8
```

```
## [1] -3.499483  3.499483
```

```r
a_sqrt_8 / sqrt(8)
```

```
## [1] -1.237254  1.237254
```

We can see here that $a = 1.237254$ is the solution.
