

WST 311

Assignment E: 5-16 March 2018

1. The scores obtained by 87 students in three subtests are given in the dataset College.txt on ClickUP.

Let $X_1 =$ Social science and history score (entry in first column);
 $X_2 =$ Verbal score (entry in second column);
 $X_3 =$ Science score (entry in third column).

The summary statistics are given in the following SAS output:

The MEANS Procedure			
Variable	N	Mean	Std Dev
social	87	526.586	76.211
verbal	87	54.690	11.227
science	87	25.126	4.807

Assume that $\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Calculate the maximum likelihood estimates for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

2. Let $\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix}$ be a vector of random variables from a multivariate normal $N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distribution with

$$\boldsymbol{\mu} = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} 9 & -1 & 2 & 0 \\ -1 & 7 & 3 & -1 \\ 2 & 3 & 13 & 3 \\ 0 & -1 & 3 & 9 \end{pmatrix}.$$

Generate a random sample of size 10 000 from this distribution and use the simulated data to calculate empirical values for the following.

- (a) $P(\mathbf{X} < \mathbf{1}_4)$.
- (b) $P(X_1 < 1, X_2 < 1, X_3 < 1)$.
- (c) $P(X_1 + X_2 > X_3 + X_4)$.
- (d) $\text{var}(X_4)$.
- (e) $\text{cov}(X_1 + X_2, X_3 + X_4)$.

3. Suppose that the heights of married couples can be described by a bivariate normal distribution. The wives have a mean height of 169.7 cm and a standard deviation of 5.1 cm. The heights of husbands have a mean of 177.8 cm and a standard deviation of 6.3 cm. The correlation between the heights of husbands and wives is 0.68. Answer the following questions by making use of a simulation.
 - (a) What is the probability that for a randomly selected couple the wife is taller than her husband?
 - (b) Consider only couples where the husband is shorter than 175cm. What is the probability that for a randomly selected couple from this group the wife is taller than her husband?
4. Work through Example A, Question 7.

SAS/IML notes for Assignment E

The document from the following website explains how to generate random samples from multivariate distributions.

http://support.sas.com/kb/26/addl/fusion26093_1_randmv.pdf.

The document is also available on ClickUP under Assignment 5.

If $\mathbf{X} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then the following syntax can be used to generate a random sample of 20 vectors from this distribution and write it into the matrix $\mathbf{X} : 20 \times 3$. The vector mu and matrix sigma must be specified.

```
proc iml; reset nolog;
mu= ...;
sigma= ...;
call randseed(0);
x=randnormal(20,mu,sig);
```

The following syntax calculates the mean of the entries in the **columns** of the matrix \mathbf{X} . If $\mathbf{X} : 20 \times 3$ then mean will be a 1×3 vector.

```
mean=x[ : ,];
```

The following syntax creates a 20×3 vector where the entries of each row is the elements of the vector mean.

```
xmean=repeat(mean,20);
```