Ec240a, Fall 2017

Professor Bryan Graham

Problem Set 1

Due: October 23rd, 2017

Problem sets are due at 5PM in the GSIs mailbox. You may work in groups, but each student should turn in their own write-up (including a printout of a narrated/commented and executed Jupyter Notebook if applicable). Please also e-mail a copy of any Jupyter Notebook to the GSI (if applicable).

1 Multivariate normal distribution

Let $\mathbf{Y} = (Y_1, \ldots, Y_K)'$ be a $K \times 1$ random vector with density function

$$f(y_1, \ldots, y_K) = (2\pi)^{-K/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{y}-\mu)' \Sigma^{-1} (\mathbf{y}-\mu)\right),$$

for Σ a symmetric positive definite $K \times K$ matrix and μ a $K \times 1$ vector. We say that $\mathbf{Y}$ is a multivariate normal random variable with mean μ and covariance Σ or

$$\mathbf{Y} \sim \mathcal{N}(\mu, \Sigma).$$

The multivariate normal distribution arises frequently in econometrics and a mastery of its basic properties is essential for both applied and theoretical work in econometrics. This problem provides an opportunity for you to review and/or learn some of these properties.

Prove the following properties of the multivariate normal distribution.

1. For
$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix}, \ \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$
with $\Sigma_{12} = \Sigma_{21}'$ show that
$$\mathbf{Y}_1 \sim \mathcal{N}(\mu_1, \Sigma_{11}).$$

2. Likewise show that
$$\mathbf{Y}_2 | \mathbf{Y}_1 = \mathbf{y}_1 \sim \mathcal{N}\left(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(\mathbf{y}_1 - \mu_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right).$$

3. If $\Sigma_{12} = 0$, show that $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are independent.

4. Let $\mathbf{Z} = \mathbf{A} + \mathbf{B}\mathbf{Y}$ for $\mathbf{A}$ and $\mathbf{B}$ non-random and $\mathbf{B}$ with full row rank, show that
$$\mathbf{Z} \sim N(\mathbf{A} + \mathbf{B}\mu, \mathbf{B}\Sigma\mathbf{B}').$$

5. Set $\mathbf{A} = 0$ and $\mathbf{B} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$. Show that for $\mathbf{Z} = \mathbf{A} + \mathbf{B}\mathbf{Y} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$ that
   (a) Z_1 and Z_2 are univariate normal random variables;

(b) their joint distribution is *not* bivariate normal;

(c) Explain.

6. Let $\{\mathbf{Y}_i\}_{i=1}^N$ be a random sample of size N drawn from the multivariate normal population described above. Show that $\sqrt{N}\left(\overline{\mathbf{Y}}-\mu\right)$ is a $\mathcal{N}(0,\Sigma)$ random variable for $\overline{\mathbf{Y}}=\frac{1}{N}\sum_{i=1}^N \mathbf{Y}_i$, the sample mean (HINT: Use independence of the $i=1,\ldots,N$ draws and your result in Problem 4 above).

7. Let $W=(\mathbf{Y}-\mu)'\Sigma^{-1}(\mathbf{Y}-\mu)$. Show that $W\sim\chi^2_K$.

8. Let $\mathbf{W}=N\cdot\left(\overline{\mathbf{Y}}-\mu\right)'\Sigma^{-1}\left(\overline{\mathbf{Y}}-\mu\right)$. Show that $\mathbf{W}\sim\chi^2_K$ (i.e., $\mathbf{W}$ is a chi-square random variable with K degrees of freedom).

9. Let $\chi_K^{2,1-\alpha}$ be the $(1-\alpha)^{th}$ quantile of the χ^2_K distribution (i.e., the number satisfying the equality $\Pr\left(\mathbf{W}\leq\chi_K^{2,1-\alpha}\right)=1-\alpha$ with $\mathbf{W}$ a chi-square random variable with K degrees of freedom). Let D be a $P\times K$ $(P\leq K)$ matrix of rank P and d a $P\times 1$ vector of constants. Consider the hypothesis

$$\begin{aligned} H_0 &: D\mu = d \\ H_1 &: D\mu \neq d. \end{aligned}$$

Maintaining H_0 derive the sampling distribution of $D\overline{\mathbf{Y}}$ as well as that of

$$\mathbf{W}=N\cdot\left(D\overline{\mathbf{Y}}-d\right)'(D\Sigma D)^{-1}\left(D\overline{\mathbf{Y}}-d\right).$$

You observe that, for the sample in hand, $\mathbf{W}>\chi_P^{2,1-\alpha}$ for $\alpha=0.05$. Assuming H_0 is true, what is the ex ante (i.e., pre-sample) probability of this event? What are you inclined to conclude after observing $\mathbf{W}$ in the sample in hand?

2 Exercises

1. [Adapted from exercise 18.1 from Goldberger (1991)]. Suppose that $\mathbf{Y}\sim\mathcal{N}(\mu,\Sigma)$ with

$$\mu=\begin{pmatrix}1\\2\\3\end{pmatrix},\ \Sigma=\begin{pmatrix}2&-1&1\\-1&5&1\\1&1&3\end{pmatrix}.$$

(a) Calculate $\mathbb{E}\left[Y_3|\,Y_1=y_1,Y_2=y_2\right]$ and $\mathbb{V}\left(Y_3|\,Y_1=y_1,Y_2=y_2\right)$

(b) Calculate $\mathbb{E}\left[Y_3|\,Y_1=y_1\right]$ and $\mathbb{V}\left(Y_3|\,Y_1=y_1\right)$

(c) Calculate $\Pr\left(-1\leq Y_3\leq 2\right)$

2. Let Y_1 and Y_0 respectively denote child and parent height. Assume that

$$Y_t\sim\mathcal{N}\left(\mu,\sigma^2\right)$$

for $t=0,1$ (so that the distribution of height is the *same* across the two generations). Let $\rho=\mathbb{C}\left(Y_1,Y_0\right)/\sqrt{\mathbb{V}\left(Y_1\right)}\sqrt{\mathbb{V}\left(Y_0\right)}$ equal the correlation between Y_0 and Y_1. Show the following:

(a) $\mathbb{E}\left[Y_1|\,Y_0=y_0\right]=(1-\rho)\,\mu+\rho y_0$

(b) Under what conditions would you expect a child's height to exceed that of their parents? The opposite. Why is this called regression to mean?

(c) Prove that, in general, $0 \leq \rho^2 \leq 1$.

3. Complete the following exercises from Hansen (2017): **2.8**, **2.9**, **2.18** (part (a) only).