556: MATHEMATICAL STATISTICS I

MULTIVARIATE EXPECTATIONS: COVARIANCE AND CORRELATION

Recall that for vector random variable $\underset{\sim}{X} = (X_1, \ldots, X_k)^\mathsf{T}$, and vector function $\underset{\sim}{g}(.)$, we have that

$$E_{f_{\underset{\sim}{X}}}[\, \underset{\sim}{g}(\underset{\sim}{X})\,] = \int \cdots \int \underset{\sim}{g}(\underset{\sim}{x}) f_{\underset{\sim}{X}}(\underset{\sim}{x})\, d\underset{\sim}{x}$$

We can consider **multivariate moments** (or **cross moments**) by choosing a particular scalar g: for integers $r_1, r_2, \ldots, r_k \geq 0$,

$$E_{f_{\underset{\sim}{X}}}[\, \underset{\sim}{g}(\underset{\sim}{X})\,] = E_{f_{\underset{\sim}{X}}}\left[X_1^{r_1} X_2^{r_2} \ldots X_k^{r_k}\right].$$

The multivariate version of generating functions can be used to compute such moments; recall that

$$M_{\underset{\sim}{X}}(\underset{\sim}{t}) = E_{f_{\underset{\sim}{X}}}\left[\exp\left\{\sum_{j=1}^{k} t_j X_j\right\}\right]$$

If $r = r_1 + r_2 + \cdots + r_k$, where each r_j is a non-negative integer, we have that

$$\frac{\partial^r}{\partial t_1^{r_1} \partial t_2^{r_2} \cdots \partial t_k^{r_k}} \left\{M_X(\underset{\sim}{t})\right\}_{\underset{\sim}{t}=\mathbf{0}} = E_{f_{\underset{\sim}{X}}}\left[X_1^{r_1} X_2^{r_2} \ldots X_k^{r_k}\right].$$

For example, if $k = 2$, we have that

$$\frac{\partial^2}{\partial t_1 \partial t_2} \left\{M_{X_1,X_2}(t_1, t_2)\right\}_{t_1=0, t_2=0} = E_{f_{X_1,X_2}}[X_1 X_2].$$

COVARIANCE AND CORRELATION

- The **covariance** of two random variables X_1 and X_2 is denoted $Cov_{f_{X_1,X_2}}[X_1, X_2]$, and is defined by
$$Cov_{f_{X_1,X_2}}[X_1, X_2] = E_{f_{X_1,X_2}}[(X_1 - \mu_1)(X_2 - \mu_2)] = E_{f_{X_1,X_2}}[X_1 X_2] - \mu_1 \mu_2$$
where $\mu_i = E_{f_{X_i}}[X_i]$ is the marginal expectation of X_i, for $i = 1, 2$.

- The **correlation** of X_1 and X_2 is denoted $Corr_{f_{X_1,X_2}}[X_1, X_2]$, and is defined by
$$Corr_{f_{X_1,X_2}}[X_1, X_2] = \frac{Cov_{f_{X_1,X_2}}[X_1, X_2]}{\sqrt{Var_{f_{X_1}}[X_1] Var_{f_{X_2}}[X_2]}}$$

NOTES:

(i) If
$$Cov_{f_{X_1,X_2}}[X_1, X_2] = Corr_{f_{X_1,X_2}}[X_1, X_2] = 0$$
then variables X_1 and X_2 are **uncorrelated**. Note that if random variables X_1 and X_2 are independent then
$$Cov_{f_{X_1,X_2}}[X_1, X_2] = E_{f_{X_1,X_2}}[X_1 X_2] - E_{f_{X_1}}[X_1] E_{f_{X_2}}[X_2] = E_{f_{X_1}}[X_1] E_{f_{X_2}}[X_2] - E_{f_{X_1}}[X_1] E_{f_{X_2}}[X_2] = 0$$
and so X_1 and X_2 are also uncorrelated (note that the converse does not hold).

(ii) For random variables X_1 and X_2, with (marginal) expectations μ_1 and μ_2 respectively, and (marginal) variances σ_1^2 and σ_2^2 respectively, if random variables Z_1 and Z_2 are defined

$$Z_1 = \frac{X_1 - \mu_1}{\sigma_1} \qquad\qquad Z_2 = \frac{X_2 - \mu_2}{\sigma_2}$$

that is, Z_1 and Z_2 are *standardized* variables. Then

$$Corr_{f_{X_1,X_2}}[X_1, X_2] = Cov_{f_{Z_1,Z_2}}[Z_1, Z_2].$$

(iii) **Extension to k variables**: covariances can only be calculated for *pairs* of random variables, but if k variables have a joint probability structure it is possible to construct a $k \times k$ *matrix*, $\mathbf{C_X}$ say, of covariance values, whose (i, j)th element is

$$Cov_{f_{X_i,X_j}}[X_i, X_j]$$

for $i, j = 1, .., k$, that captures the complete covariance structure in the joint distribution. If $i \neq j$, then

$$Cov_{f_{X_j,X_i}}[X_j, X_i] = Cov_{f_{X_i,X_j}}[X_i, X_j]$$

so $\mathbf{C_X}$ is *symmetric*, and if $i = j$,

$$Cov_{f_{X_i,X_i}}[X_i, X_i] \equiv Var_{f_{X_i}}[X_i]$$

The matrix $\mathbf{C_X}$ is referred to as the **variance-covariance** matrix, and we can write

$$\mathbf{C_X} = Var_{f_{\underset{\sim}{X}}}[\underset{\sim}{X}].$$

(iv) If $\underset{\sim}{X}$ is a $k \times 1$ vector random variable with variance-covariance matrix $\mathbf{C_X}$, let $\mathbf{A}$ be a $d \times k$ matrix. Then $\underset{\sim}{Y} = \mathbf{A}\underset{\sim}{X}$ is a $d \times 1$ vector random variable, and

$$Var_{f_{\underset{\sim}{Y}}}[\underset{\sim}{Y}] = Var_{f_{\underset{\sim}{X}}}[\mathbf{A}\underset{\sim}{X}] = \mathbf{A}Var_{f_{\underset{\sim}{X}}}[\underset{\sim}{X}]\mathbf{A}^{\mathsf{T}}$$

so that

$$\mathbf{C_Y} = \mathbf{A}\mathbf{C_X}\mathbf{A}^{\mathsf{T}}$$

is the $d \times d$ variance-covariance matrix for $\underset{\sim}{Y}$.

(v) If random variable X is defined by $X = a_1X_1 + a_2X_2 + \ldots + a_kX_k$, for random variables $X_1, ..., X_k$ and constants $a_1, ..., a_k$, then

$$\begin{aligned} E_{f_X}[X] &= \sum_{i=1}^{k} a_i E_{f_{X_i}}[X_i] \\ Var_{f_X}[X] &= \sum_{i=1}^{k} a_i^2 Var_{f_{X_i}}[X_i] + 2\sum_{i=1}^{k}\sum_{j=1}^{i-1} a_i a_j Cov_{f_{X_i,X_j}}[X_i, X_j] \end{aligned}$$

(vi) Combining the results above when $k = 2$, and defining standardized variables Z_1 and Z_2 as in (ii), we have

$$\begin{aligned} 0 \leq Var_{f_{Z_1,Z_2}}[Z_1 \pm Z_2] &= Var_{f_{Z_1}}[Z_1] + Var_{f_{Z_2}}[Z_2] \pm 2Cov_{f_{Z_1,Z_2}}[Z_1, Z_2] \\ &= 1 + 1 \pm 2Corr_{f_{X_1,X_2}}[X_1, X_2] = 2(1 \pm Corr_{f_{X_1,X_2}}[X_1, X_2]) \end{aligned}$$

and hence

$$-1 \leq Corr_{f_{X_1,X_2}}[X_1, X_2] \leq 1.$$