2.7. MULTIVARIATE EXPECTATIONS AND COVARIANCE

For random variables $X_1, ..., X_k$ with range $\mathbb{X}^{(k)}$ with mass/density function $f_{X_1,...,X_k}$, the **expectation** of $g(X_1, ..., X_k)$ is defined in the discrete and continuous cases by

$$\mathrm{E}_{f_{X_1,...,X_k}}[\ g(X_1, ..., X_k)\ ] = \begin{cases} \displaystyle\sum_{x_1} ... \sum_{x_k} g(x_1, ..., x_k) f_{X_1,...,X_k}(x_1, ..., x_k) \\ \displaystyle\int_{\mathbb{X}_1} ... \int_{\mathbb{X}_k} g(x_1, ..., x_k) f_{X_1,...,X_k}(x_1, ..., x_k)\ dx_1 ... dx_k \end{cases}$$

PROPERTIES

(i) Let g and h be real-valued functions and let a and b be constants. Then, if $f_{\mathbf{X}} \equiv f_{X_1,...,X_k}$,

$$\mathrm{E}_{f_{\mathbf{X}}}[\ ag(X_1, ..., X_k) + bh(X_1, ..., X_k)\ ] = a\mathrm{E}_{f_{\mathbf{X}}}[\ g(X_1, ..., X_k)\ ] + b\mathrm{E}_{f_{\mathbf{X}}}[\ h(X_1, ..., X_k)\ ].$$

(ii) Let $X_1, ... X_k$ be **independent** random variables with mass functions/pdfs $f_{X_1}, ..., f_{X_k}$ respectively. Let $g_1, ..., g_k$ be scalar functions of $X_1, ..., X_k$ respectively (that is, g_i is a function of X_i *only* for $i = 1, ..., k$). If $g(X_1, ..., X_k) = g_1(X_1) ... g_k(X_k)$, then

$$\mathrm{E}_{f_{\mathbf{X}}}[\ g(X_1, ..., X_k)\ ] \ = \ \prod_{i=1}^{k} \mathrm{E}_{f_{X_i}}[\ g_i(X_i)\ ]$$

where $\mathrm{E}_{f_{X_i}}[\ g_i(X_i)\ ]$ is the marginal expectation of $g_i(X_i)$ with respect to f_{X_i}.

COVARIANCE AND CORRELATION

The **covariance** of two random variables X_1 and X_2 is denoted $\mathrm{Cov}_{f_{X_1,X_2}}[X_1, X_2]$, and is defined by

$$\mathrm{Cov}_{f_{X_1,X_2}}[X_1, X_2] = \mathrm{E}_{f_{X_1,X_2}}[(X_1 - \mu_1)(X_2 - \mu_2)] = \mathrm{E}_{f_{X_1,X_2}}[X_1 X_2] - \mu_1 \mu_2$$

where $\mu_i = \mathrm{E}_{f_{X_i}}[X_i]$ is the marginal expectation of X_i, for $i = 1, 2$, and where $\mathrm{E}_{f_{X_1,X_2}}[X_1 X_2]$ is the expectation of function $g(x_1, x_2) = x_1 x_2$ with respect to the joint distribution f_{X_1,X_2}.

The **correlation** of X_1 and X_2 is denoted $\mathrm{Corr}_{f_{X_1,X_2}}[X_1, X_2]$, and is defined by

$$\mathrm{Corr}_{f_{X_1,X_2}}[X_1, X_2] = \frac{\mathrm{Cov}_{f_{X_1,X_2}}[X_1, X_2]}{\sqrt{\mathrm{Var}_{f_{X_1}}[X_1]\ \mathrm{Var}_{f_{X_2}}[X_2]}}$$

If $\mathrm{Cov}_{f_{X_1,X_2}}[X_1, X_2] = \mathrm{Corr}_{f_{X_1,X_2}}[X_1, X_2] = 0$. then variables X_1 and X_2 are **uncorrelated**. Note that if random variables X_1 and X_2 are independent then

$$\mathrm{Cov}_{f_{X_1,X_2}}[X_1, X_2] = \mathrm{E}_{f_{X_1,X_2}}[X_1 X_2] - \mathrm{E}_{f_{X_1}}[X_1]\mathrm{E}_{f_{X_2}}[X_2] = \mathrm{E}_{f_{X_1}}[X_1]\mathrm{E}_{f_{X_2}}[X_2] - \mathrm{E}_{f_{X_1}}[X_1]\mathrm{E}_{f_{X_2}}[X_2] = 0$$

and so X_1 and X_2 are also uncorrelated (the converse does not hold).

NOTES:
(i) 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

$$Z_1 = (X_1 - \mu_1)/\sigma_1 \qquad Z_2 = (X_2 - \mu_2)/\sigma_2$$

that is, Z_1 and Z_2 are *standardized* variables. Then $\mathrm{Corr}_{f_{X_1,X_2}}[X_1, X_2] = \mathrm{Cov}_{f_{Z_1,Z_2}}[Z_1, Z_2]$.

(ii) 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*, $\mathbf{C}$ say, of covariance values, whose (i, j)th element is

$$\mathrm{Cov}_{f_{X_i,X_j}}[X_i, X_j]$$

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

$$\mathrm{Cov}_{f_{X_j,X_i}}[X_j, X_i] = \mathrm{Cov}_{f_{X_i,X_j}}[X_i, X_j]$$

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

$$\mathrm{Cov}_{f_{X_i,X_i}}[X_i, X_i] \equiv \mathrm{Var}_{f_{X_i}}[X_i]$$

The matrix $\mathbf{C}$ is referred to as the **variance-covariance** matrix.

(iii) If random variable X is defined by

$$X = \sum_{i=1}^{k} a_i X_i$$

for random variables $X_1, ..., X_k$ and constants $a_1, ..., a_k$, then

$$\begin{aligned} \mathrm{E}_{f_X}[X] &= \sum_{i=1}^{k} a_i \mathrm{E}_{f_{X_i}}[X_i] \\ \mathrm{Var}_{f_X}[X] &= \sum_{i=1}^{k} a_i^2 \mathrm{Var}_{f_{X_i}}[X_i] + 2\sum_{i=1}^{k}\sum_{j=1}^{i-1} a_i a_j \mathrm{Cov}_{f_{X_i,X_j}}[X_i, X_j] \end{aligned}$$

(iv) Combining (i) and (iii) when $k = 2$, and defining standardized variables Z_1 and Z_2,

$$\begin{aligned} 0 \leq \mathrm{Var}_{f_{Z_1,Z_2}}[Z_1 \pm Z_2] &= \mathrm{Var}_{f_{Z_1}}[Z_1] + \mathrm{Var}_{f_{Z_2}}[Z_2] \pm 2\ \mathrm{Cov}_{f_{Z_1,Z_2}}[Z_1, Z_2] \\ &= 1 + 1 \pm 2\ \mathrm{Corr}_{f_{X_1,X_2}}[X_1, X_2] \\ &= 2(1 \pm \mathrm{Corr}_{f_{X_1,X_2}}[X_1, X_2]) \end{aligned}$$

and hence $-1 \leq \mathrm{Corr}_{f_{X_1,X_2}}[X_1, X_2] \leq 1$.

2.8. SUMS OF RANDOM VARIABLES - THE CONVOLUTION FORMULA

Suppose that X_1 and X_2 be continuous random variables taking values on $\mathbb{X}_1$ and $\mathbb{X}_2$ espectively with joint density function f_{X_1,X_2}. Then, if random variable Y is defined by $Y = X_1 + X_2$, the marginal density of Y is given by calculated as follows; consider the multivariate transformation

$$\begin{array}{rlll} \mathbf{g}: & (X_1, X_2) & \longrightarrow (Y, Z) & \\ & (x_1, x_2) & \longmapsto (y, z) & \text{where } y = x_1 + x_2, z = x_1 \end{array}$$

so that the inverse transformation is given by $(X_1, X_2) = (Z, Y - Z)$. Then by the multivariate transformation theorem,

$$f_{Y,Z}(y,z) = f_{X_1,X_2}(z, y-z)|J(y,z)| \qquad \text{for } (y,z) \in B = \{(y,z)|(y,z) = \mathbf{g}(x_1,x_2), x_1 \in \mathbb{X}_1, x_2 \in \mathbb{X}_2\}.$$

$J(y,z)$ is the determinant of a 2×2 matrix of partial derivatives of the inverse functions ${g_1}^{-1}(t_1,t_2) = t_2$, ${g_1}^{-1}(t_1,t_2) = t_1 - t_2$, which simply reduces to give $J(y,z) = -1$. Therefore

$$f_{Y,Z}(y,z) = f_{X_1,X_2}(z, y-z)$$

and so by the usual marginalization calculation

$$f_Y(y) = \int_{\mathbb{Z}} f_{Y,Z}(y,z)\ dz = \int_{\mathbb{X}_1} f_{X_1,X_2}(x_1, y - x_1)\ dx_1$$

If X_1 and X_2 are **independent**, then $f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$ so if $Y = X_1 + X_2$, then Y has pdf given by

$$f_Y(y) = \int_{\mathbb{X}_1} f_{X_1}(x_1) f_{X_2}(y - x_1)\ dx_1$$

2.9. ORDER STATISTICS

For k random variables $X_1, ..., X_k$, the **order statistics**, $Y_1, ..., Y_k$, are defined by

$$Y_i = X_{(i)} \text{ - “the } i\text{th smallest value in } X_1, ..., X_k\text{”}$$

for $i = 1, ..., k$, so that $Y_1 = X_{(1)} = \text{Min}\{X_1, ..., X_k\}$, $Y_k = X_{(k)} = \text{Max}\{X_1, ..., X_k\}$.
For k independent, identically distributed random variables $X_1, ..., X_k$, with marginal density function f_X, there are two main results to consider;

RESULT 1 The joint density function of the order statistics $Y_1, ..., Y_k$ is given by

$$f_{Y_1,...,Y_k}(y_1, ..., y_k) = k! f_X(y_1)...f_X(y_k) \qquad y_1 < ... < y_k$$

RESULT 2 The marginal pdf of the jth order statistic Y_j for $j = 1, ..., k$ has the form

$$f_{Y_j}(y_j) = \frac{k!}{(j-1)!(k-j)!} \{F_X(y_j)\}^{j-1} \{1 - F_X(y_j)\}^{k-j} f_X(y_j)$$

Special Cases: To derive the marginal pdf of Y_k, first consider the marginal cdf of Y_k;

$$\begin{aligned} F_{Y_k}(y_k) &= \mathrm{P}[Y_k \le y_k] = \mathrm{P}[\max\{X_1, ..., X_k\} \le y_k] = \mathrm{P}[X_1 \le y_k, X_2 \le y_k, ..., X_k \le y_k] \\ &= \prod_{i=1}^{k} \mathrm{P}[X_i \le y_k] = \prod_{i=1}^{k} \{F_X(y_k)\} = \{F_X(y_k)\}^k \\ \Longrightarrow f_{Y_k}(y_k) &= k\{F_X(y_k)\}^{k-1} f_X(y_k) \qquad \text{(by differentiation)} \end{aligned}$$

By a similar calculation, we can find the marginal pdf/cdf for Y_1,

$$F_{Y_1}(y_1) = 1 - \{1 - F_X(y_1)\}^k \qquad f_{Y_1}(y_1) = k\{1 - F_X(y_1)\}^{k-1} f_X(y_1)$$

2.10. GENERATING FUNCTIONS

For random variable X with mass/density function f_X, the **moment generating function**, or **mgf**, of X, M_X, is defined by

$$M_X(t) = \mathrm{E}_{f_X}[e^{tX}]$$

if this expectation exists for all values of $t \in (-h, h)$ for some $h > 0$. It can be shown that if X_1 and X_2 are random variables taking values on $\mathbb{X}$ with mass/density functions f_{X_1} and f_{X_2}, and mgfs M_{X_1} and M_{X_2} respectively, then

$$f_{X_1}(x) \equiv f_{X_2}(x),\ x \in \mathbb{X} \quad \Longleftrightarrow \quad M_{X_1}(t) \equiv M_{X_2}(t),\ t \in (-h, h)$$

As seen previously, we have that the rth derivative of M_X evaluated at 0, $M_X^{(r)}(0)$, gives the rth moment of X

$$M_X^{(r)}(0) = \sum_{x \in \mathbb{X}} x^r f_X(x) = \mathrm{E}_{f_X}[X^r] \qquad \Longleftrightarrow \qquad M_X(t) = 1 + \sum_{r=1}^{\infty} \frac{t^r}{r!} \mathrm{E}_{f_X}[X^r]$$

Also, from the general result for expectations of functions of random variableS we have that if random variable Y is defined in terms of X by $Y = aX + b$,

$$\mathrm{E}_{f_Y}[e^{tY}] \equiv \mathrm{E}_{f_X}[e^{t(aX+b)}] \Longrightarrow M_Y(t) = \mathrm{E}_{f_X}[e^{t(aX+b)}] = e^{bt}\mathrm{E}_{f_X}[e^{atX}] = e^{bt}M_X(at).$$

THEOREM

Let $X_1, ..., X_k$ be independent random variables with mgfs $M_{X_1}, ..., M_{X_k}$ respectively. Then if random variable Y is defined by $Y = X_1 + ... + X_k$,

$$M_Y(t) = \prod_{i=1}^{k} M_{X_i}(t)$$

Proof Using the general result for expectations of functions of independent random variables,

$$M_Y(t) = \mathrm{E}_{f_Y}[e^{tY}] = \mathrm{E}_{f_{X_1,...,X_k}}[e^{t(X_1+...+X_k)}] = \prod_{i=1}^{k} \mathrm{E}_{f_{X_i}}[e^{tX_i}] = \prod_{i=1}^{k} M_{X_i}(t).$$

Special Case If $X_1, ..., X_k$ are **identically distributed**, then $M_{X_i}(t) \equiv M_X(t)$, say, for all i, so

$$M_Y(t) = \prod_{i=1}^{k} M_X(t) = \{M_X(t)\}^k$$