# 4 Multivariate Distributions

## 4.1 Joint Distributions

1. **Joint distribution function**: If $X_1, \ldots, X_n$ are random variables, the joint (cumulative) distribution function is

$$F(x_1, x_2, \ldots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n).$$

We sometimes use the notation $F_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$ when we need to distinguish the joint distribution function of $X_1, \ldots, X_n$ from other distribution functions.

2. **Bivariate distribution function**: For two random variable $X$ and $Y$ the joint distribution functions is

$$F(x, y) = P(X \leq x, Y \leq y).$$

We sometimes use the notation $F_{X,Y}(x, y)$ for the joint distribution function of $X$ and $Y$.

   (a) $F(-\infty, y) = \lim_{x\to-\infty} F(x, y) = 0$,
   $F(x, -\infty) = \lim_{y\to-\infty} F(x, y) = 0$,
   $F(\infty, \infty) = \lim_{x\to\infty,y\to\infty} F(x, y) = 1$.

   (b) $P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = F(x_2, y_2) - F(x_1, y_2) - [F(x_2, y_1) - F(x_1, y_1)]$.

   (c) Right continuous is $x$: $\lim_{h\downarrow 0} F(x + h, y) = F(x, y)$,
   Right continuous in $y$: $\lim_{h\downarrow 0} F(x, y + h) = F(x, y)$.

3. **Marginal distribution functions**: If $F_{X,Y}$ is the joint distribution function of $X$ and $Y$ then the marginal distribution functions are given by

$$F_X(x) = \lim_{y\to\infty} F_{X,Y}(x, y) = F_{X,Y}(x, \infty),$$
$$F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x, y) = F_{X,Y}(\infty, y).$$

4. **Discrete bivariate distributions**: for discrete random variables $X$ and $Y$.

   (a) Joint mass function: $f(x, y) = f_{X,Y}(x, y) = P(X = x, Y = y)$.

   (b) Probabilities: $P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = \sum_{x_1<x\leq x_2} \sum_{y_1<y\leq y_2} f(x, y)$.

   (c) Marginal mass functions: $f_X(x) = \sum_y f_{X,Y}(x, y)$,
   $f_Y(y) = \sum_x f_{X,Y}(x, y)$.

5. **Continuous bivariate distributions**: for jointly continuous random variables $X$ and $Y$.

   (a) Joint density function: is a positive real-valued function $f$ such that

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u, v)dudv \quad \text{for each } x, y \in \mathbb{R}.$$

   (b) Probabilities: $P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = \int_{y_1}^{y_2} \int_{x_1}^{x_2} f(x, y)dxdy$.

   (c) Marginal density functions: $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy$,
   $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dx$,

6. **Expectation of a function of two variables**: If $g$ is a well-behaved, real-valued function of two variables ($g : \mathbb{R}^2 \longrightarrow \mathbb{R}$) and $X$ and $Y$ are random variable with joint mass/density function $f_{X,Y}$ then

$$E[g(X,Y)] = \begin{cases} \sum_y \sum_x g(x,y) f_{X,Y}(x,y), & \text{discrete case,} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy, & \text{continuous case.} \end{cases}$$

7. **Multivariate generalizations**: for random variables $X_1, \ldots, X_n$ and for $j = 1, \ldots, n$.

   (a) Marginal distribution: $F_{X_j}(x_j) = F_{X_1,\ldots,X_n}(\infty, \ldots, \infty, x_j, \infty, \ldots, \infty)$

   (b) Marginal mass/density

$$f_{X_j}(x_j) = \begin{cases} \sum_{x_1} \cdots \sum_{x_{j-1}} \sum_{x_{j+1}} \cdots \sum_{x_n} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n), & \text{discrete case,} \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) dx_1 \ldots dx_{j-1} dx_{j+1} \ldots dx_n, & \text{continuous case.} \end{cases}$$

   (c) Expectation of a function of $n$ variables:

$$E[g(X_1, \ldots, X_n)] = \begin{cases} \sum_{x_1} \cdots \sum_{x_n} g(x_1, \ldots, x_n) f_{X_1,\ldots,X_n}(x_1, \ldots, x_n), & \text{discrete,} \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n) f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) dx_1 \ldots dx_n, & \text{continuous.} \end{cases}$$

## 4.2 Conditional Distributions

1. **Discrete conditional distributions**: $X$ and $Y$ discrete random variables with $P(X = x) > 0$.

   (a) Conditional distribution function of $Y$ given $X = x$: $F_{Y|X}(y|x) = P(Y \leq y | X = x)$.

   (b) Conditional mass function of $Y$ given $X = x$: $f_{Y|X}(y|x) = P(Y = y | X = x) = f_{X,Y}(x,y)/f_X(x)$.

   (c) Relationship between distribution and mass functions: $F_{Y|X}(y|x) = \sum_{y_i \leq y} f_{Y|X}(y_i|x)$.

2. **Continuous conditional distributions**: $X$ and $Y$ jointly continuous random variables with $f_X(x) > 0$.

   (a) Conditional distribution function of $Y$ given $X = x$: $F_{Y|X}(y|x) = \int_{-\infty}^{y} (f_{X,Y}(x,v)/f_X(x)) dv$.

   (b) Conditional density function of $Y$ given $X = x$: $f_{Y|X}(y|x) = f_{X,Y}(x,y)/f_X(x)$.

3. **Conditional, joint and marginal densities**: given $f_X(x) > 0$ and (for (d)) $f_Y(y) > 0$.

   (a) Conditional mass/density

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} \frac{f_{X,Y}(x,y)}{\sum_y f_{X,Y}(x,y)}, & \text{discrete case,} \\ \frac{f_{X,Y}(x,y)}{\int_{-\infty}^{\infty} f_{X,Y}(x,y) dy}, & \text{continuous case.} \end{cases}$$

   (b) Joint mass/density: $f_{X,Y}(x,y) = f_{Y|X}(y|x) f_X(x)$.

   (c) Marginal mass/density

$$f_Y(y) = \begin{cases} \sum_x f_{Y|X}(y|x) f_X(x), & \text{discrete case,} \\ \int_{-\infty}^{\infty} f_{Y|X}(y|x) f_X(x) dx, & \text{continuous case.} \end{cases}$$

   (d) Reverse conditioning

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(x)} = \frac{f_X(x)}{f_Y(y)} f_{Y|X}(y|x).$$

4. **Conditional expectation**: $X$ and $Y$ random variables.

   (a) Condition expectation of $Y$ given $X$: define

   $$\psi(x) = E(Y|X=x) = \begin{cases} \sum_y y f_{Y|X}(y|x), & \text{discrete case,} \\ \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy, & \text{continuous case.} \end{cases}$$

   The conditional expectation of $Y$ given $X$ is $E(Y|X) = \psi(X)$ (a random variable).

   (b) Condition expectation of $g(Y)$ given $X$: If $g$ is a well-behaved, real-valued function, define

   $$h(x) = E(g(Y)|X=x) = \begin{cases} \sum_y g(y) f_{Y|X}(y|x), & \text{discrete case,} \\ \int_{-\infty}^{\infty} g(y) f_{Y|X}(y|x) dy, & \text{continuous case.} \end{cases}$$

   The conditional expectation of $g(Y)$ given $X$ is $E(g(Y)|X) = h(X)$ (a random variable).

   (c) Law of iterated expectations:

   $$E[\psi(X)] = E[E(Y|X)] = E(Y).$$

   Useful consequence,

   $$E(Y) = \begin{cases} \sum_x E(Y|X=x) f_X(x), & \text{discrete case,} \\ \int_{-\infty}^{\infty} E(Y|X=x) f_X(x) dx, & \text{continuous case.} \end{cases}$$

5. **Conditional variance**: for random variables $X$ and $Y$, define

   $$\omega(x) = \text{Var}(Y|X=x) = \begin{cases} \sum_y [y - E(Y|X=x)]^2 f_{Y|X}(y|x), & \text{discrete case,} \\ \int_{-\infty}^{\infty} [y - E(Y|X=x)]^2 f_{Y|X}(y|x) dy, & \text{continuous case.} \end{cases}$$

   The conditional variance of $Y$ given $X$ is $\text{Var}(Y|X) = \omega(X) = E(Y^2|X) - [E(Y|X)]^2$. The conditional variance is a random variable. Using the law of iterated expectations, we can show that $\text{Var}(Y) = \text{Var}[E(Y|X)] + E[\text{Var}(Y|X)]$.

## 4.3 Dependence

1. **Independence of random variables** The random variables $X_1,\ X_2,\dots,\ X_n$ are (mutually) independent if and only if the events $\{X_1 \le x_1\},\ \{X_2 \le x_2\},\dots,\ \{X_n \le x_n\}$ are independent for all choices of $x_1, x_2, \dots, x_n$.

   If $X_1, X_2, \dots, X_n$ independent, then for all $x_1, x_2, \dots, x_n$:

   (a) $F_{X_1,\dots,X_n}(x_1,\dots,x_n) = F_{X_1}(x_1)F_{X_2}(x_2)\dots F_{X_n}(x_n)$,

   (b) $f_{X_1,\dots,X_n}(x_1,\dots,x_n) = f_{X_1}(x_1)f_{X_2}(x_2)\dots f_{X_n}(x_n)$,

   (c) $E(X_1X_2\dots X_n) = E(X_1)E(X_2)\dots E(X_n)$,

   (d) $g_1(X_1), g_2(X_2), \dots, g_n(X_n)$ are independent for real-valued functions $g_1, g_2, \dots, g_n$.

2. **Covariance function**: for random variables $X$ and $Y$,

$$\mathrm{Cov}(X,Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y).$$

   (a) Symmetry: $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$.

   (b) With constant multipliers: $\mathrm{Cov}(aX, bY) = ab\mathrm{Cov}(X,Y)$.

   (c) Bilinearity: $\mathrm{Cov}(X_1+X_2, Y_1+Y_2) = \mathrm{Cov}(X_1,Y_1)+\mathrm{Cov}(X_1,Y_2)+\mathrm{Cov}(X_2,Y_1)+\mathrm{Cov}(X_2,Y_2)$.

   (d) Variance:
$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{Cov}(X,X),\\ \mathrm{Var}(X+Y) &= \mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y),\\ \mathrm{Var}(X-Y) &= \mathrm{Var}(X)+\mathrm{Var}(Y)-2\mathrm{Cov}(X,Y).\end{aligned}$$

   (e) If $X$ and $Y$ are independent, $\mathrm{Cov}(X,Y) = 0$.

3. **Correlation coefficient**: for random variables $X$ and $Y$,

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}.$$

   Correlation is scaled covariance, $|\mathrm{Corr}(X,Y)| \le 1$. Independent implies uncorrelated but the reverse implication does not hold.

## 4.4 Joint Moments

If $X$ and $Y$ are random variables with joint mass/density function $f_{X,Y}$ then the $(r,s)^{\mathrm{th}}$ joint moment is

$$m_{r,s} = E(X^rY^s) = \begin{cases} \sum_y \sum_x x^r y^s f_{X,Y}(x,y), & \text{discrete case,}\\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^r y^s f_{X,Y}(x,y)dxdy, & \text{continuous case.}\end{cases}$$

The $(r,s)^{\mathrm{th}}$ joint central moment is

$$\mu_{r,s} = E[(X-E(X))^r(Y-E(Y))^s] = \begin{cases} \sum_y \sum_x [(x-\mu_X)^r(y-\mu_Y)^s] f_{X,Y}(x,y), & \text{discrete case,}\\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [(x-\mu_X)^r(y-\mu_Y)^s] f_{X,Y}(x,y)dxdy, & \text{continuous case.}\end{cases}$$

1. **Properties of joint moments**

   (a) $r^{\mathrm{th}}$ moment for $X$: $m_{r,0} = E(X^r)$.

   (b) $r^{\mathrm{th}}$ central moment for $X$: $\mu_{r,0} = E[(X-\mu_X)^r]$.

   (c) Covariance: $\mu_{1,1} = E[(X-E(X))(Y-E(Y))] = \mathrm{Cov}(X,Y)$.

2. **Joint moment generating function**

$$M_{X,Y}(t,u) = E(e^{tX+uY}) = \begin{cases} \sum_y \sum_x e^{tx+uy} f_{X,Y}(x,y), & \text{discrete case,} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{tx+uy} f_{X,Y}(x,y)dxdy, & \text{continuous case.} \end{cases}$$

   (a) Coefficient of $t^r u^s$:

$$\frac{1}{r!s!}E(X^rY^s) = \frac{1}{r!s!}m_{r,s}$$

.

   (b) Moment generating function for marginal: $M_X(t) = E(e^{tX}) = M_{X,Y}(t,0)$, $M_Y(t) = E(e^{uY}) = M_{X,Y}(0,u)$.

   (c) Derivatives at zero:

$$M_{X,Y}^{(r,s)}(0,0) = \frac{d^{r+s}}{dt^r du^s}M_{X,Y}(t,u)\bigg|_{t=0,u=0} = E(X^rY^s) = m_{r,s}.$$

   (d) If $X$ and $Y$ independent: $M_{X,Y}(t,u) = M_X(t)M_Y(u)$.

3. **Conditional moment generating function**: define

$$\nu(u,x) = M_{Y|X}(u|x) = E(e^{uY}|X=x) = \begin{cases} \sum_y e^{uy} f_{Y|X}(y|x), & \text{discrete case,} \\ \int_{-\infty}^{\infty} e^{uy} f_{Y|X}(y|x)dy, & \text{continuous case.} \end{cases}$$

The conditional moment generating function of $Y$ given $X$ is $M_{Y|X}(u|X) = \nu(u,X) = E(e^{uY}|X)$. This is a conditional expectation so it is a random variable. We can calculate joint moment generating function and moment generating function for marginal $Y$ from the conditional moment generating function,

$$\begin{aligned} M_{X,Y}(t,u) &= E(e^{tX+uY}) = E[e^{tX}M_{Y|X}(u|X)], \\ M_Y(u) &= M_{X,Y}(0,u) = E[M_{Y|X}(u|X)]. \end{aligned}$$

4. **Joint cumulants**: let $K_{X,Y}(t,u) = \log M_{X,Y}(t,u)$, then we define the $(r,s)^{\text{th}}$ joint cumulant $\kappa_{r,s}$ as the coefficient of $(t^ru^s)/(r!s!)$ in the expansion of $K_{X,Y}$. Thus, $\kappa_{1,1} = \text{Cov}(X,Y)$.

5. **Multivariate generalization**: for random variables $X_1,\dots,X_n$ with joint mass/density function $f_{X_1,\dots,X_n}$.

   (a) Joint moments:

$$\begin{aligned} m_{r_1,\dots,r_n} &= E(X_1^{r_1}\dots X_n^{r_n}) \\ &= \begin{cases} \sum_{x_1}\dots\sum_{x_n} x_1^{r_1}\dots x_n^{r_n} f_{X_1,\dots,X_n}(x_1,\dots,x_n), & \text{discrete case,} \\ \int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} x_1^{r_1}\dots x_n^{r_n} f_{X_1,\dots,X_n}(x_1,\dots,x_n)dx_1\dots dx_n, & \text{continuous case.} \end{cases} \end{aligned}$$

   (b) Joint central moments: $\mu_{r_1,\dots,r_n} = E[(X_1 - E(X_1))^{r_1}\dots(X_n - E(X_n))^{r_n}]$

   (c) Joint moment generating function: $M_{X_1,\dots,X_n}(t_1,\dots,t_n) = E(e^{t_1X_1+\dots+t_nX_n})$. The coefficient of $t_1^{r_1}\dots t_n^{r_n}$ in the expansion of $M_{X_1,\dots,X_n}$ is $E(X_1^{r_1}\dots X_n^{r_n})/(r_1!\dots r_n!)$.

   (d) Joint cumulant generating function: $K_{X_1,\dots,X_n}(t_1,\dots,t_n) = \log(M_{X_1,\dots,X_n}(t_1,\dots,t_n))$. The $(r_1,\dots,r_n)^{\text{th}}$ joint cumulant is defined as the coefficient of $(t_1^{r_1}\dots t_n^{r_n})/(r_1!\dots r_n!)$ in the expansion of $K_{X_1,\dots,X_n}$.

   (e) Independence: if $X_1,\dots,X_n$ are independent then $M_{X_1,\dots,X_n}(t_1,\dots,t_n) = M_{X_1}(t_1)\dots M_{X_n}(t_n)$.