Joint distributions

Notes: Below X and Y are assumed to be continuous random variables. This case is, by far, the most important case. Analogous formulas, with sums replacing integrals and p.m.f.'s instead of p.d.f.'s, hold for the case when X and Y are discrete r.v.'s. Appropriate analogs also hold for mixed cases (e.g., X discrete, Y continuous), and for the more general case of n random variables $X_1, \dots, X_n$.

- **Joint cumulative distribution function (joint c.d.f.):**

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

- **Joint density (joint p.d.f.):** Given a joint c.d.f. $F(x,y)$, the associated joint p.d.f. is given by $f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$.

  In general, a joint density function is any (integrable) function $f(x,y)$ satisfying the properties

$$f(x,y) \ge 0, \qquad \iint f(x,y)dxdy = 1.$$

  Usually, $f(x,y)$ will be given by an explicit formula, along with a *range* (a region in the xy-plane) on which this formula holds. *In the general formulas below, if a range of integration is not explicitly given, the integrals are to be taken over the range in which the density function is defined.*

- **Marginal densities:** The ordinary (one-variable) densities of X and Y, denoted by f_X and f_Y. The marginal densities can be computed from the joint density $f(x,y)$ via the formulas

$$f_X(x) = \int f(x,y)dy, \quad f_Y(y) = \int f(x,y)dx.$$

- **Conditional densities:** The conditional density of X given that $Y = y$, and of Y given that $X = x$ are defined by

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

  provided the denominators, $f_Y(y)$, resp., $f_X(y)$, are non-zero. Regarded as a function of x (with y fixed), the conditional density $f_{X|Y}(x|y)$ satisfies the properties of an ordinary (one-variable) density function; the same goes for $f_{Y|X}(y|x)$.

- **Probabilities via joint densities:** Given a region B in the xy-plane, the probability that (X,Y) falls into this region is given by the double integral of $f(x,y)$ over this region:

$$P((X,Y) \in B) = \iint_B f(x,y)dxdy$$

- **Expectations via joint densities:** Given a function of x and y (e.g., $g(x,y) = xy$, or $g(x,y) = x$, etc.),

$$E(g(X,Y)) = \iint g(x,y)f(x,y)dxdy.$$

- **Independence:** X and Y are called independent if the joint p.d.f. is the product of the individual p.d.f.'s, i.e., if

  $$f(x,y) = f_X(x)f_Y(y) \quad \text{for all } x,\ y.$$

  Note: While the individual (marginal) densities f_X and f_Y can always be computed from the joint density $f(x,y)$, only for independent r.v.'s can one go backwards, i.e., obtain the joint density from the marginal densities.

  Properties of independent random variables: If X and Y are independent, then:

  - The product formula holds for probabilities of the form P(some condition on X, some condition on Y) (where the comma denotes "and"). For example, $P(X \le 2, Y \le 3) = P(X \le 2)P(Y \le 3)$.
  - The expectation of the product of X and Y is the product of the individual expectations, $E(XY) = E(X)E(Y)$. More generally, this product formula holds for any expectation of a function X times a function of Y. For example, $E(X^2Y^3) = E(X^2)E(Y^3)$.

- **Sums of independent r.v.'s, general formula:** If X and Y are independent r.v.'s with densities f_X and f_Y, then the density f_Z of the sum $Z = X + Y$ is given by the "convolution formula":

  $$f_Z(z) = \int f_X(x)f_Y(z-x)dx = \int f_Y(y)f_X(z-y)dy.$$

- **Sums of independent r.v.'s, special cases:**

  - **Normal distribution:** If X_i $(i = 1,\dots,n)$ are independent normal $N(\mu_i, \sigma_i^2)$, then, for any constants c_i:
    * $\sum_{i=1}^n c_iX_i$ is normal $N(\mu, \sigma^2)$, where $\mu = \sum_{i=1}^n c_i\mu_i$ and $\sigma^2 = \sum_{i=1}^n c_i^2\sigma_i^2$.

    In particular, if X_1 and X_2 are independent and normal $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively, then:
    * $X_1 + X_2$ is normal $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.
    * $X_1 - X_2$ is normal $N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.

    (Note that the variances, σ_1^2 and σ_2^2, add up in both cases.)
  - **Poisson distribution:** If X_i $(i = 1,\dots,n)$ are independent Poisson r.v.'s with parameters λ_i, then:
    * $\sum_{i=1}^n X_i$ is Poison with parameter $\lambda = \sum_{i=1}^n \lambda_i$.
  - **Binomial distribution:** If X_i $(i = 1,\dots,r)$ are independent binomial r.v.'s with parameters n_i and p, then:
    * $\sum_{i=1}^r X_i$ is binomial with parameters n and p, where $n = \sum_{i=1}^r n_i$.

    (Note that the second parameter, p, here must be the same for all X_i.)