Soc 30005: Statistical Methods of Research 2

Notes on Probability

I. Probability for Discrete Variables

Joint Probability:

$$\Pr(Y = y, X = x),\ x \in \{0,1,...\},\ y \in \{0,1,...\}$$

Marginal Probability

$$\begin{aligned}\Pr(Y = y) &= \sum_x \Pr(Y = y, X = x)\\ \Pr(X = x) &= \sum_y \Pr(Y = y, X = x)\end{aligned} \tag{1}$$

Conditional Probability

$$\Pr(Y = y \mid X = x) = \frac{\Pr(Y = y, X = x)}{\Pr(X = x)} \tag{2}$$

$$\Pr(X = x \mid Y = y) = \frac{\Pr(Y = y, X = x)}{\Pr(Y = y)}$$

Decompositions

Joint Probability: From (2),

$$\begin{aligned}\Pr(Y = y, X = x) &= \Pr(Y = y \mid X = x)\Pr(X = x)\\ &= \Pr(X = x \mid Y = y)\Pr(Y = y)\end{aligned}$$

Marginal Probability: From (1) and (2)

$$\Pr(Y = y) = \sum_x \Pr(Y = y \mid X = x)\Pr(X = x)$$

$$\Pr(X = x) = \sum_y \Pr(Y = y \mid X = x)\Pr(X = x)$$

Bayes Theorem

$$\begin{aligned}\Pr(Y=y \mid X=x) &= \frac{\Pr(Y=y, X=x)}{\Pr(X=x)} \\ &= \frac{\Pr(X=x \mid Y=y)\Pr(Y=y)}{\Pr(X=x)} \\ &= \frac{\Pr(X=x \mid Y=y)\Pr(Y=y)}{\sum_y \Pr(X=x \mid Y=y)\Pr(Y=y)}\end{aligned}$$

Independence

If X and Y are independent, $\Pr(Y=y \mid X=x) = \Pr(Y=y)$.

Therefore,

$$\begin{aligned}\Pr(Y=y, X=x) &= \Pr(Y=y \mid X=x)\Pr(X=x) \\ &= \Pr(Y=y)\Pr(X=x)\end{aligned}$$

II. Probability for Continuous Variables

Henceforth, we shall refer to Prob(Y=y) as f(y).

Marginal Probability

$$\Pr(a < X < b) \approx \sum_{i=1}^{n} f(x_i)\Delta$$

$$\Pr(a < X < b) = \lim_{\substack{n\to\infty \\ \Delta\to 0}} \left(\sum_{i=1}^{n} f(x_i)\Delta \right) = \int_a^b f(x)dx$$

[where $x_i = a + i\,\Delta$; $\Delta = (b-a)/n$].

We must have $f(x) \geq 0$, $\int_{-\infty}^{\infty} f(x)dx = 1$.

Probability density function

In the continuous case, $f(x)$ is called the *probability density* function.

Joint Probability

$$\Pr(a < X < b, c < Y < d) = \int_c^d \int_a^b f(x,y)dxdy$$

Note: $f(x) = \int\int f(x,y)dy; \quad f(y) = \int\int f(x,y)dx$ (3)

Joint probability density

$f(x,y)$ is the *joint density* function

Conditional Probability

$$\Pr(c < Y < d \mid a < X < b) = \frac{\int_c^d \int_a^b f(x,y)dxdy}{\int_a^b f(x)dx} = \frac{\int_c^d \int_a^b f(x,y)dxdy}{\int_a^b \int_{-\infty}^{\infty} f(x,y)dydx}$$

Conditional probability density

$f(y \mid x) = \dfrac{f(x,y)}{f(x)}$ is the *conditional density* function. (4)

Decompositions

Joint Probability: From (4),

$$f(x,y) = f(y \mid x)f(x) = f(x \mid y)f(y)$$

Marginal Probability: From (3) and (4)

$$f(x) = \int f(x,y)dy = \int f(y \mid x)f(x)dy$$

$$f(y) = \int f(x,y)dy = \int f(x \mid y)f(y)dx$$

Bayes Theorem

$$f(y \mid x) = \frac{f(x,y)}{f(x)} = \frac{f(x \mid y)f(y)}{f(x)} = \frac{f(x \mid y)f(y)}{\int f(x \mid y)f(y)dy}.$$

Independence

If X and Y are independent, $f(x,y) = f(x)f(y)$.

III. Expectation

Discrete case: $E(X) = \sum_{i=1}^{n} x_i f(x_i)$

Continuous case: $E(X) = \int x_i f(x_i)$

Sums of Expectations

$E(X+Y) = E(X) + E(Y)$; therefore $E(kX) = kE(X)$ (5)

Conditional Expectation

$$E(Y \mid X) = \begin{Bmatrix} \sum y_i f(y_i \mid x) & \textit{discrete case} \\ \int y_i f(y_i \mid x)\, dy & \textit{continuous case} \end{Bmatrix}$$

Independence

If $X \perp Y$, $E(Y \mid X) = E(Y)$

Proof (discrete case): Recall that if $X \perp Y$, $f(y \mid x) = f(y)$. Then $E(Y \mid X) = \sum y_i f(y_i \mid x) = \sum y_i f(y_i) = E(Y)$.

III. Variance

$Var(X) = E(X - \mu_x)^2$, where $\mu_x = E(X)$.

Note $E(X - \mu_x)^2 = E(X^2) - \mu_x^2$

IV. Covariance

$Cov(X,Y) = E(X - \mu_x)(Y - \mu_y)$

Note if $X \perp Y$, $Cov(X,Y) = Cov(X,Y) = E(X - \mu_x)E(Y - \mu_y) = 0$ (6)

V. Variance of a Sum

$$\begin{aligned} Var(X+Y) &= E[(X+Y) - (\mu_x \mu_y)]^2 \\ &= E[(X - \mu_x) + (Y - \mu)]^2 \\ &= E[(X - \mu_x)^2 + (Y - \mu)^2 + 2(X - \mu_x)(Y - \mu)] \\ &= Var(X) + Var(Y) + 2Cov(X,Y) \end{aligned}$$

From (6), if $X \perp Y$, $Var(X+Y) = Var(X) + Var(Y)$

Common notation

$$Var(X) = \sigma_x^2; \;\; Var(Y) = \sigma_y^2, \;\; Cov(X,Y) = \sigma_{xy}$$

Or

$$Var(X) = \sigma_{xx}; \;\; Var(Y) = \sigma_{yy}, \;\; Cov(X,Y) = \sigma_{xy}$$