Discrete Conditional Distributions

- Suppose that X and Y are jointly discrete random variables with probability mass function

$$p_{X,Y}(x,y) = \mathsf{P}(X = x, Y = y)$$

- If we know the value of Y then we can find the conditional probability

$$\mathsf{P}(X = x \mid Y = y) = \frac{\mathsf{P}(X = x, Y = y)}{\mathsf{P}(Y = y)} = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

- We can find this probability for each possible value of x.

- This gives us the conditional Probability Mass Function of X given $Y = y$.

Definition 4.8
Suppose that X and Y are jointly discrete random variables with joint probability mass function $p_{X,Y}(x,y)$ and that the marginal probability mass function of Y is $p_Y(y)$. Let y be a possible value of Y with $p_Y(y) > 0$ then the conditional probability mass function of X given $Y = y$ is

$$p_{X|Y}(x \mid y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}.$$

Definition 4.9
Suppose that the conditional probability mass function of X given $Y = y$ is $p_{X|Y}(x \mid y)$, the conditional cumulative distribution of X given $Y = y$ is

$$F_{X|Y}(x \mid y) = \mathsf{P}(X \leqslant x \mid Y = y) = \sum_{t \leqslant x} p_{X|Y}(t, y).$$

Definition 4.10

Suppose that X and Y are discrete random variables then the conditional mean and conditional variance of X given $Y = y$ are

$$\begin{aligned} \mathsf{E}[X \mid Y = y] &= \sum_x x p_{X|Y}(x \mid y) \\ \mathsf{Var}(X \mid Y = y) &= \sum_x \Big(x - \mathsf{E}[X \mid Y = y]\Big)^2 p_{X|Y}(x \mid y) \\ &= \sum_x x^2 p_{X|Y}(x \mid y) - \left(\sum_x x p_{X|Y}(x \mid y)\right)^2 \end{aligned}$$

Theorem 4.17

If X and Y are two independent discrete random variables then the conditional probability mass function of X given $Y = y$ is the marginal probability mass function of X.

Continuous Conditional Distributions

- If Y is a continuous random variable then $\mathsf{P}(Y = y) = 0$ so we cannot divide by this.

- Suppose that h is extremely small then

$$\mathsf{P}\Big(Y \in [y, y + h)\Big) = \int_y^{y+h} f_Y(v)\, dv \approx h f_Y(y).$$

- Similarly we can say that

$$\mathsf{P}\Big(X \in A, Y \in [y, y+h)\Big) = \int_A \int_y^{y+h} f_{X,Y}(x, v)\, dv\, dx \approx \int_A h f_{X,Y}(x, y)\, dx.$$

Hence for an extremely small h we have

$$\begin{aligned}
\mathsf{P}\Big(X \in A \mid Y \in [y, y+h)\Big) &= \frac{\mathsf{P}\Big(X \in A, Y \in [y, y+h)\Big)}{\mathsf{P}\Big(Y \in [y, y+h)\Big)} \\
&\approx \frac{\int_A h f_{X,Y}(x,y)\,dx}{h f_Y(y)} \\
&= \int_A \frac{f_{X,Y}(x,y)}{f_Y(y)}\,dx
\end{aligned}$$

- Since h is extremely small, we can think of this as the conditional probability $\mathsf{P}(X \in A \mid Y = y)$.

- Comparing the above to

$$\mathsf{P}(X \in A) = \int_A f_X(x)\,dx$$

  We can think of the integrand as the probability density function for X conditional on $Y = y$.

Definition 4.11

Suppose that X and Y are two continuous random variables then the conditional probability density function of X given $Y = y$ is

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

Definition 4.12

Suppose the X and Y are two continuous random variables the the conditional cumulative distribution function of X given $Y = y$ is

$$F_{X|Y}(x \mid y) = \mathsf{P}(X \leqslant x \mid Y = y) = \int_{-\infty}^{x} f_{X|Y}(u, y)\, du.$$

Definition 4.13

Suppose that X and Y are continuous random variables then the conditional mean and conditional variance of X given $Y = y$ are

$$\begin{aligned}
\mathsf{E}[X \mid Y = y] &= \int_{-\infty}^{\infty} x f_{X|Y}(x \mid y)\, dx \\
\mathsf{Var}(X \mid Y = y) &= \int_{-\infty}^{\infty} \Big(x - \mathsf{E}[X \mid Y = y]\Big)^2 f_{X|Y}(x \mid y)\, dx \\
&= \int_{-\infty}^{\infty} x^2 f_{X|Y}(x \mid y)\, dx - \left(\int_{-\infty}^{\infty} x f_{X|Y}(x \mid y)\, dx\right)^2
\end{aligned}$$

Theorem 4.18

If X and Y are two independent continuous random variables then the conditional probability density function of X given $Y = y$ is the marginal probability density function of X.