Conditional Distributions

- Review conditional pmf/pdf, conditional mean, and conditional variance.
- Example 3-5 from `ConditionalDistributions.pdf`.

Covariance and Correlation Coefficient

- pp 5-8 from `Independence_and_Covariance.pdf`.
- Key concepts: **Covariance** and **Correlation**

Covariance of X and Y

$$\sigma_{XY} = \mathsf{Cov}(X,Y) = \mathbb{E}\big[(X-\mu_X)(Y-\mu_Y)\big] = \mathbb{E}(XY) - \mu_X\mu_Y$$

- Connection with Variance: $\mathsf{Cov}(X,X) = \mathsf{Var}(X)$;
- Symmetric: $\mathsf{Cov}(X,Y) = \mathsf{Cov}(Y,X)$;
- Linearity with **scale** change, but invariant with **location** change

$$\mathsf{Cov}(aX+b,Y) = a\mathsf{Cov}(X,Y), \quad \mathsf{Cov}(X+Y,W) = \mathsf{Cov}(X,W)+\mathsf{Cov}(Y,W).$$

$$\begin{aligned} \mathsf{Cov}(aX+bY, cX+dY) &= ac\mathsf{Var}(X) + bd\mathsf{Var}(Y) \\ &\quad +(ad+bc)\mathsf{Cov}(X,Y) \\ \mathsf{Var}(aX+bY) &= a^2\mathsf{Var}(X) + 2ab\mathsf{Cov}(X,Y) + b^2\mathsf{Var}(Y). \end{aligned}$$

Cauchy-Schwarz Inequality

$$|\text{Cov}(X,Y)| \leq \sqrt{\text{Var}(X)\text{Var}(Y)}$$

Proof:

$$\begin{aligned}\text{Var}(aX - bY) \geq 0 \quad &\Longrightarrow \quad a^2\text{Var}(X) - 2ab\text{Cov}(X,Y) + b^2\text{Var}(Y) \geq 0, \\ &\Longrightarrow \quad a^2\sigma_X^2 + b^2\sigma_Y^2 \geq 2ab \times \sigma_{XY}, \qquad (1)\end{aligned}$$

for any a, b values. In particular, (1) holds true with

$$a = \frac{\sigma_Y}{\sigma_{XY}} = \frac{\sqrt{\text{Var}(Y)}}{\text{Cov}(X,Y)}, \quad b = \frac{\sigma_X}{\sigma_{XY}} = \frac{\sqrt{\text{Var}(X)}}{\text{Cov}(X,Y)},$$

which proves the CS inequality.

Correlation Coefficient

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}(X)\mathsf{Var}(Y)}}.$$

- $-1 \le \rho_{XY} \le 1$, due to the CS inequality.
- $\rho_{XY} = \pm 1$ **if and only if** X and Y are linear functions of one other.
- The magnitude of ρ_{XY} reflects **only the linear dependence** between X and Y. So it is possible that $Y = g(X)$ where g is a one-to-one map (i.e., X totally determines Y), but ρ_{XY} is small.
- $\rho_{XY} = 0$ **if and only if** $\mathsf{Cov}(X,Y) = 0$, then we say X and Y are uncorrelated.
- If X and Y are independent, then $\rho_{XY} = 0$, but the reverse doesn't hold.

Mean and Variance of Sum of Random Variables

Let $X_1, \ldots, X_n$ be n random variables and $a_0, a_1, \ldots, a_n$ be $(n+1)$ constants. Define $U = a_0 + a_1 X_1 + \cdots + a_n X_n$.

$$\mathbb{E}U = \mathbb{E}\big(a_0 + a_1 X_1 + \cdots + a_n X_n\big) = a_0 + a_1 \mathbb{E}X_1 + \cdots + a_n \mathbb{E}X_n$$

$$\begin{aligned}
\mathsf{Var}(U) &= \mathsf{Var}\big(a_0 + a_1 X_1 + \cdots + a_n X_n\big) \\
&= \sum_{i=1}^{n} a_i^2 \mathsf{Var}(X_i) + \sum_{i \neq j} a_i a_j \mathsf{Cov}(X_i, X_j) \\
&= \sum_{i=1}^{n} a_i^2 \mathsf{Var}(X_i) + 2 \sum_{i<j} a_i a_j \mathsf{Cov}(X_i, X_j)
\end{aligned}$$

If X_i's are **uncorrelated**, i.e., $\mathsf{Cov}(X_i, X_j) = 0$ for all $i \neq j$, then variance of the sum is equal to the sum of variances.