

Conditional Probability and Distribution Functions: Lecture IV

Charles B. Moss

September 9, 2026

- 1 Conditional Probability and Independence
 - Bayes Theory
- 2 Useful Distribution Functions
 - Normal Distributions
 - Uniform Distributions
 - Gamma Distribution
- 3 Transformation of Random Variables
 - Inverse Hyperbolic Sine
- 4 Triangular Distribution Function

Conditional Probability and Independence

- In order to define the concept of a conditional probability it is necessary to define joint and marginal probabilities.
 - The joint probability is the probability of a particular combination of two or more random variables.
 - Taking the role of two die as an example, the probability of rolling a 4 on one die and a 6 on the other die is $1/36$.
 - There are 36 possible outcomes of the two die $\{1, 1\}$, $\{1, 2\}$, $\dots$ $\{2, 1\}$, $\{2, 2\}$, $\dots$ $\{6, 6\}$.
 - Therefore the probability of a $\{4, 6\}$ given that the die are fair is $1/36$.

- The marginal probability is the probability one of the random variables irrespective of the outcome of the other variable.
 - Going back to the die example, there are six different rolls of the die where the value of the first die is 4

$$\{4, 1\}, \{4, 2\}, \{4, 3\}, \{4, 4\}, \{4, 5\}, \{4, 6\} \quad (1)$$

- Hence, again assume that the die are fair the marginal probability of $x_1 = 4$ is

$$\begin{aligned} P[x_1 = 4] &= P[\{4, 1\}] + P[\{4, 2\}] + P[\{4, 3\}] + \\ &\quad P[\{4, 4\}] + P[\{4, 5\}] + P[\{4, 6\}] \\ &= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} \\ &= \frac{6}{36} = \frac{1}{6} \end{aligned} \quad (2)$$

- The conditional probability is then the probability of one event, such as the probability that the first die is a 4, given that the value of another random variable is known, such as the fact that the value of the second die roll is equal to 6. In the forgoing example, the case of the fair die, this value is $1/6$.
- **Definition 2.10** *p-18* Given that A and B are sets defined on C the Axioms of Conditional Probability are
 - $P[A|B] \geq 0$ for all A ,
 - $P[A|A] = 1$,
 - If $\{A_i \cap B_i\}$ are mutually exclusive events

$$P[A_1 \cup A_2 \cup \dots \cup A_n | B] = P[A_1|B] + P[A_2|B] + \dots + P[A_n|B] \quad (3)$$

- If $B \supset H$, $B \supset G$, and $P[G] \neq 0$, then

$$\frac{P[H|B]}{P[G|B]} = \frac{P[H]}{P[G]} \quad (4)$$

- Discussion of Definition 2.10.
 - The first 3 conditions follow the general axioms of probability theory.
 - The final condition states that the relative probability of a conditional event and marginal distribution functions are the same.
- Given this construction the conditional probability can be derived as

$$P[A|B] = \frac{P[A \cap B]}{P[B]} \quad (5)$$

- In our discussion of the role of the die

$$P[x_1 = 4, x_2 = 6] = \frac{P[x_1 = 4 \text{ and } x_2 = 6]}{P[x_2 = 6]} = \frac{1/36}{1/6} = \frac{1}{6} \quad (6)$$

Bayes Theory

- Given that the events $A_1, A_2, \dots, A_n$ are mutually exclusive events such that $P[A_1 \cup A_2 \cup \dots A_n] = 1$ the conditional probability can then be extended to Bayes Theorem

$$P[A_i|E] = \frac{P[E|A_i] P[A_i]}{\sum_{j=1}^n P[E|A_j] P[A_j]} \quad (7)$$

Given that

$$P[A|B] = \frac{P[A \cap B]}{P[B]} \Rightarrow P[B] P[A|B] = P[A \cap B] \quad (8)$$

- Substituting this result into the previous expression

$$P[A_i|E] = \frac{P[A_i \cap E]}{\sum_{j=1}^n P[A_j \cap E]} \quad (9)$$

- This expression unifies the simple expression in Equation 5. Specifically, following the Axioms of Probability (Condition 3).

$$P[A_1] + P[A_2] + \cdots P[A_n] = P[A_1 \cup A_2 \cup \cdots A_n] \quad (10)$$

- Thus, the denominator of Equation 7 becomes

$$P[A_1 \cap E] + P[A_2 \cap E] + \cdots P[A_n \cap E] = P[(A_1 \cup A_2 \cup \cdots A_n) \cap E] \quad (11)$$

- Given that $A_1, A_2, \cdots A_n$ are exhaustive, or $A_1 \cup A_2 \cup \cdots A_n = C$ the entire sample space since $P[A_1 \cup A_2 \cup \cdots A_n] = 1$

$$P[(A_1 \cup A_2 \cup \cdots A_n) \cap E] = P[C \cap E] = P[E] \quad (12)$$

with the last equality since $E \subset C$.

- Returning to the numerator in Equation 7

$$P[E] = \sum_{i=1}^n P[E|A_i] P[A_i] \quad (13)$$

- Definition 2.11** *p-20* Two events A and B are independent if $P[A] = P[A|B]$.
 - In die example

$$P[x_1 = 4] = P[x_1 = 4|x_2 = 6] = \frac{1}{6} \quad (14)$$

- Independence of three random variables.
 - $P[A \cap B] = P[A] P[B]$
 - $P[A \cap C] = P[A] P[C]$
 - $P[B \cap C] = P[B] P[C]$
 - $P[A \cap B \cap C] = P[A] P[B] P[C]$

Some Useful Distribution Functions

- Univariate normal distribution

$$f(x|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right] \quad (15)$$

- Multivariate normal distribution

$$f(x|\mu, \Sigma) = \frac{1}{\sqrt{2\pi}} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2} (x - \mu)' \Sigma^{-1} (x - \mu)\right) \quad (16)$$
$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix} \quad \rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

Uniform Distributions

- Univariate uniform distribution

$$f(x|a, b) = \begin{cases} \frac{1}{b-a} & \text{if } x \in [a, b] \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

$$\begin{aligned} \int_{-\infty}^{\infty} f(x|a, b) dx &= \int_{-\infty}^a 0 dx + \int_a^b \frac{1}{b-a} dx + \int_b^{\infty} 0 dx \\ &= \left(\frac{x}{b-a} \right) \Big|_a^b = \frac{b}{b-a} - \frac{a}{b-a} = \frac{b-a}{b-a} = 1 \end{aligned} \quad (18)$$

- Multivariate uniform distribution

$$f(x_1, x_2) = \begin{cases} 1 & \text{if } 0 \leq x_1 \leq 1 \text{ and } 0 \leq x_2 \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

- Integral of the multivariate uniform

$$\begin{aligned}\int_0^1 \int_0^1 f(x_1, x_2) dx_1 dx_2 &= \int_0^1 \left((x|_0^1) \right) dx_2 = \int_0^1 (1 - 0) dx_2 \\ &= \int_0^1 dx_2 = (x|_0^1 = 1 - 0 = 1\end{aligned}\tag{20}$$

- To examine the conditional properties of the bivariate uniform distribution, we start by deriving the marginal distribution of x_1 . This marginal distribution is derived by integrating out x_2

$$f_1(x_1) = \int_0^1 f(x_1, x_2) dx_2 = \int_0^1 dx_2 = (x_2|_0^1 = 1 \tag{21}$$

- The conditional distribution for x_2 given the value of x_1 is then written as

$$f(x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)} = \frac{1}{1} = 1 \quad (22)$$

Therefore we conclude that x_1 and x_2 are independent since

$$f_2(x_2) = f(x_2|x_1) = 1 \quad (23)$$

Univariate Gamma

$$f(x|\alpha, \beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right) & \text{for } 0 < x < \infty \text{ and } \beta > 0 \\ 0 & \text{otherwise} \end{cases} \quad (24)$$

Transformation of Random Variables

- Another alternative is to create a new distribution by transforming one random variable into another with a known distribution.
 - Nested probability functions

$$F(x^*) = \int_{-\infty}^{x^*} f(x) dx \Leftrightarrow P[x \leq x^*] \quad (25)$$

$$\begin{aligned} z = \phi(x) &\Rightarrow x = \phi^{-1}(z) \\ g(z) &= f(\phi^{-1}(z)) \left| \frac{d\phi^{-1}(z)}{dz} \right| \end{aligned} \quad (26)$$

Inverse Hyperbolic Sine distribution

$$\begin{aligned} e_t = e_t(\epsilon_t, \theta) &= \frac{\ln \left(\theta \epsilon_t + \left[(\theta \epsilon_t)^2 + 1 \right]^{\frac{1}{2}} \right)}{\theta} \\ g(\epsilon_t | \mu, \sigma^2, \theta) &= \frac{1}{\sigma \sqrt{2\pi}} \exp \left[-\frac{1}{2} \frac{(e_t(\epsilon_t, \theta) - \mu)^2}{\sigma^2} \right] \\ &\quad \times (e_t^2(\epsilon_t, \theta) \theta^2 + 1) \end{aligned} \quad (27)$$

Triangular Distribution Function

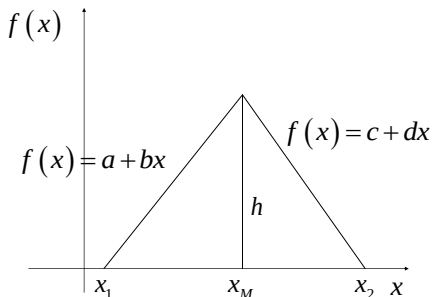


Figure: Triangular Distribution Function

$$f(x) = \begin{cases} 0 & \text{if } x < x_1 \\ a + bx & \text{if } x_1 \leq x < x_M \\ c + dx & \text{if } x_M \leq x < x_2 \\ 0 & \text{if } x \geq x_2 \end{cases} \quad (28)$$
$$\frac{1}{2}(x_M - x_1)h + \frac{1}{2}(x_2 - x_M)h = 1$$
$$h = \frac{2}{(x_M - x_1) + (x_2 - x_M)} = \frac{2}{x_2 - x_1}$$

$$b = \frac{\frac{2}{(x_2 - x_1)}}{(x_M - x_1)} = \frac{2}{(x_M - x_1)(x_2 - x_1)} \quad (29)$$

$$f(x_1) = a + bx_1 = 0 \Rightarrow a = -bx_1 \quad (30)$$

$$d = -\frac{\frac{2}{(x_2 - x_1)}}{(x_2 - x_M)} = -\frac{2}{(x_2 - x_M)(x_2 - x_1)} \quad (31)$$

$$f(x_2) = c + dx_2 = 0 \Rightarrow c = -dx_2 \quad (32)$$

Readings

- Moss (2010) *Risk, Uncertainty and the Agricultural Firm* Sections 2.5.
- Moss (2014) *Mathematical Statistics for Applied Econometrics* Chapter 4.