Random Variables and Probability Distributions: Lecture VI

Charles B. Moss

August 31, 2026

Caribou

Ptarmigan

Reading Review

- Consider a function defined as

$$f\left(x\right)=\begin{cases}0\ ,\ \mathrm{x}\leq-1\\ 1+x\ ,\ -1<\mathrm{x}\leq 0\\ 1-x\ ,\ 0<\mathrm{x}\leq 1\\ 0\ ,\ \mathrm{x}>1\end{cases}\tag{1}$$

  Is this function a valid distribution function? What are the "rules?"

- Featherstone, Allen M., Charles B. Moss, Timothy G. Baker, and Paul V. Preckel. 1988. The Theoretical Effects of Farm Policy on Optimal Leverage and the Probability of Equity Loss. *American Journal of Agricultural Economics* 70(3), 572-579. [DOI: 10.2307/1241495] (Google Scholar – 140, Web of Science – 40)

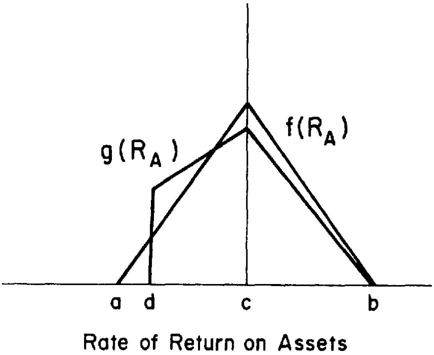

Figure 2. A mean- and mode-preserving truncation of a triangular distribution

1. Conditional Probability and Independence
    - Axioms of Conditional Probability
2. Basic Concept of Random Variables
    - Binomial Conditional Probabilities
    - Uncorrelated Discrete Normal
    - Uncorrelated Normal Conditional Probabilities
    - Correlated Discrete Normal
    - Correlated Normal Conditional Probabilities
3. Univariate Continuous Random Variables
4. Useful Univariate Distributions

Conditional Probability and Independence

- In order to define the concept of a conditional probability it is necessary to discuss joint probabilities and marginal probabilities.
  - A joint probability is the probability of two random events. For example, consider drawing two cards from the deck of cards. There are 52x51=2,652 different combinations of the first two cards from the deck.
  - The marginal probability is overall probability of a single event or the probability of drawing a given card.
  - The conditional probability of an event is the probability of that event given that some other event has occurred.
    - In the textbook, what is the probability of the die being a one if you know that the face number is odd? (1/3).
    - However, note that if you know that the role of the die is a one, that the probability of the role being odd is 1.

Axioms of Conditional Probability

- Axioms of Conditional Probability:
  1. $P(A|B) \geq 0$ for any event A.
  2. $P(A|B) = 1$ for any event $A \supset B$.
  3. If $\{ A_i \cap B_i \}$ $i = 1, 2, \ldots$ are mutually exclusive, then
  $$P(A_1 \cup A_2 \cup \ldots) = P(A_1|B) + P(A_2|B) + \ldots \tag{2}$$
  4. If $B \supset H$, $B \supset G$ and $P(G) \neq 0$ then
  $$\frac{P(H|B)}{P(G|B)} = \frac{P(H)}{P(G)} \tag{3}$$

- Theorem 2.4.1: $P(A|B) = P(A \cap B) / P(B)$ for any pair of events A and B such that $P(B) > 0$.
- Theorem 2.4.2: (*Bayes Theorem*) Let Events $A_1, A_2, \ldots A_n$ be mutually exclusive events such that $P(A_1 \cup A_2 \cup \cdots A_n) = 1$ and $P(A_i) > 0$ for each i. Let E be an arbitrary event such that $P(E) > 0$. Then

$$P(A_i|E) = \frac{P(E|A) P(A_i)}{\displaystyle\sum_{j=1}^{n} P(E|A_j) P(A_j)} \tag{4}$$

- Another manifestation of this theorem is from the joint distribution function:

$$P(E, A_i) = P(E \cap A_i) = P(E|A_i)\, P(A_i) \tag{5}$$

- The bottom equality reduces the marginal probability of event E

$$P(E) = \sum_{i=1}^{n} P(E|A_i)\, P(A_i) \tag{6}$$

- This yields a friendlier version of Bayes theorem based on the ratio between the joint and marginal distribution function:

$$P(A_i|E) = \frac{P(E, A_i)}{P(E)} \tag{7}$$

- Statistical independence is when the probability of one random variable is independent of the probability of another random variable.
  - Definition 2.4.1: Events A and B are said to be independent if $P(A) = P(A|B)$
  - Definition 2.4.2: Events A, B, and C are said to be mutually independent if the following equalities hold:
    - $P(A \cap B) = P(A)\, P(B)$
    - $P(A \cap C) = P(A)\, P(C)$
    - $P(B \cap C) = P(B)\, P(C)$
    - $P(A \cap B \cap C) = P(A)\, P(B)\, P(C)$

Basic Concept of Random Variables

- Definition 1.4.1: A random variable is a function from a sample space S into the real numbers.
- In this way a random variable is an abstraction. We assumed that there was a random variable defined on some sample space like flipping a coin. The flipping of the coin is an outcome in an abstract space

$$S = \{s_1, s_2, \cdots s_n\} \tag{8}$$

We then define a numeric value to this set of random variables

$$\begin{array}{c} X : S \rightarrow R_+^1 \\ x_i = X(s_i) \end{array} \text{ or } \begin{array}{c} X(\omega) : \Omega \rightarrow R_+^1 \\ x_i = X(\omega_i) \end{array} \tag{9}$$

The probability function (or measure) is then defined based on that random variable:

- Definitions of a Random Variable
  - Definition 3.1.1. A *random variable* is a variable that takes values according to a certain probability.
  - Definition 3.1.2. A *random variable* is a real-valued function defined over a sample space.
- Discrete Random Variables
  - Definition 3.2.1. A *discrete random variable* is a variable that takes a countable number of real numbers with certain probability.
  - Definition 3.2.2. A *bivariate discrete random variable* is a variable that takes a countable number of points on the plane with certain probability.

Table: Binomial Probability

x	y 0	1	2	3	4	5	Marginal Probability
0	0.0131	0.0436	0.0581	0.0387	0.0129	0.0017	0.1681
1	0.0280	0.0934	0.1245	0.0830	0.0277	0.0037	0.3602
2	0.0240	0.0800	0.1067	0.0711	0.0237	0.0032	0.3087
3	0.0103	0.0343	0.0457	0.0305	0.0102	0.0014	0.1323
4	0.0022	0.0073	0.0098	0.0065	0.0022	0.0003	0.0284
5	0.0002	0.0006	0.0008	0.0006	0.0002	0.0000	0.0024
Marginal Probability	0.0778	0.2592	0.3456	0.2304	0.0768	0.0102	

- In a bivariate distribution, the marginal distribution is the distribution of one variable unconditioned on the outcome of the other variable.

$$P[X = x_i] = \sum_{j=1}^{m} P[X = x_i, Y = y_j] \tag{11}$$

- Applying Bayes Theorem

$$P[X = x_i | Y = y_j] = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} \tag{12}$$

- Definition 3.2.3. Discrete random variables are said to be independent if the event $X = x_i$, and the event $Y = y_j$ are independent for all i, j . That is to say, $P(X = x_i, Y = y_j) = P(X = x_i)\, P(Y = y_j)$.

Table: Binomial Conditional Probabilities

X	$P[X,Y=2]$	$P[Y=2]$	$P[X\vert Y=2]$	$P[X]$
0	0.0581	0.3456	0.1681	0.1681
1	0.1245	0.3456	0.3602	0.3602
2	0.1067	0.3456	0.3087	0.3087
3	0.0457	0.3456	0.1323	0.1323
4	0.0098	0.3456	0.0284	0.0284
6	0.0008	0.3456	0.0024	0.0024

Note that Column 5 in Table 2 is the same as the marginal probabilities (i.e., Column 8) in Table 1. Next, we consider the discrete form of the uncorrelated normal distribution as Presented in Table 3. Again, computing the conditional distribution of X such that $Y = 2$ yields the results in Table 4. Finally, we consider a discrete form of the correlated normal distribution in Table 5.

Table: Uncorrelated Discrete Normal

	y						Marginal
x	0	1	2	3	4	5	Probability
0	0.005	0.018	0.028	0.005	0.003	0.000	0.06
1	0.023	0.055	0.088	0.070	0.010	0.000	0.25
2	0.025	0.068	0.148	0.105	0.028	0.003	0.38
3	0.010	0.063	0.100	0.045	0.010	0.003	0.23
4	0.003	0.030	0.033	0.015	0.000	0.003	0.08
5	0.000	0.000	0.008	0.003	0.000	0.000	0.01
Marginal Probability	0.07	0.23	0.40	0.24	0.05	0.01	

Table: Uncorrelated Normal Conditional Probabilities

X	$P[X,Y=2]$	$P[Y=2]$	$P[X \vert Y=2]$	$P[X]$
0	0.0275	0.4025	0.0683	0.0575
1	0.0875	0.4025	0.2174	0.2450
2	0.1475	0.4025	0.3665	0.3750
3	0.1000	0.4025	0.2484	0.2300
4	0.0325	0.4025	0.0807	0.0825
5	0.0075	0.4025	0.0186	0.0100

Table: Correlated Discrete Normal

x	y 0	1	2	3	4	5	Marginal Probability
0	0.068	0.025	0.000	0.000	0.000	0.000	0.093
1	0.028	0.115	0.078	0.003	0.000	0.000	0.223
2	0.000	0.065	0.200	0.060	0.000	0.000	0.325
3	0.000	0.000	0.088	0.143	0.020	0.000	0.250
4	0.000	0.000	0.003	0.033	0.043	0.008	0.085
5	0.000	0.000	0.000	0.003	0.010	0.013	0.025
Marginal Probability	0.095	0.205	0.368	0.240	0.073	0.020	

Table: Correlated Normal Conditional Probabilities

X	$P[X,Y=2]$	$P[Y=2]$	$P[X\vert Y=2]$	$P[X]$
0	0.0000	0.4025	0.0000	0.0925
1	0.0775	0.4025	0.1925	0.2225
2	0.2000	0.4025	0.4969	0.3250
3	0.0875	0.4025	0.2174	0.2500
4	0.0025	0.4025	0.0062	0.0850
5	0.0000	0.4025	0.0000	0.0250

- Theorem 3.2.1. Discrete Random variables X and Y with the probability distribution given in Table 3.1 are independent if and only if every row is proportional to any other row, or, equivalently, every column is proportional to any other column.
- Multivariate Random Variables
- Definition 3.2.4. A *T-variate random variable* is a variable that takes a countable number of points on the T-dimensional Euclidean space with certain probabilities.

Univariate Continuous Random Variables

- Definition 3.3.1. If there is a nonnegative function $f(x)$ defined over the whole line such that

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x)dx \tag{13}$$

  for any x_1 and x_2 satisfying $x_1 \leq x_2$, then X is a continuous random variable and $f(x)$ is called its density function.

- By axiom 2, the total area under the density function must equal 1

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{14}$$

- The simplest example of a continuous random variable is the uniform distribution

$$f(x) = \begin{cases} 1 \text{ if } 0 \leq x \leq 1 \\ 0 \text{ otherwise} \end{cases} \tag{15}$$

It is obvious that

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)dx &= \int_{-\infty}^{0} f(x)dx + \int_{0}^{1} f(x)dx + \int_{1}^{\infty} f(x)dx \\ &= \int_{0}^{1} f(x)dx = \int_{0}^{1} dx = \left(x|_{0}^{1} + C = (1-0) + C \right. \end{aligned} \tag{16}$$

- Definition 3.3.2. Let X have density $f(x)$. The *conditional density* of X given $a \leq X \leq b$, denoted by $f(x|a \leq X \leq b)$, is defined by

$$\begin{aligned} f\left(x|a \leq X \leq b\right) &= \frac{f(x)}{\int_a^b f(x)dx} \text{for } a \leq x \leq b, \\ &= 0 \text{ otherwise} \end{aligned} \tag{17}$$

- Definition 3.3.3. Let X have the density $f(x)$ and let S be a subset of the real line such that $P\left(X \in S\right) > 0$. Then the conditional density of X given $X \in S$, denoted by $f\left(x|S\right)$, is defined by

$$\begin{aligned} f\left(x\,|S\,\right) &= \tfrac{f(x)}{P(X \in S)} \text{ for x } \in \text{S} \\ &= 0 \text{ otherwise} \end{aligned} \tag{18}$$

An Almost Simple Example

- Let's start with a quadratic function defined over the range $x \in [0,2]$

$$f(x)\,1-(x-1)^2=-x^2+2x \tag{19}$$

- We can determine the constant that will make this a valid density function by integrating

$$\begin{aligned} k &= \int_0^2 \left(-x^2+2x\right)dx = -\int_0^2 x^2 dx + 2\int_0^2 x dx \\ &= -\frac{1}{3}\,x^3\Big|_0^2 + x^2\Big|_0^2 = -\frac{8}{3}+4=\frac{4}{3} \end{aligned} \tag{20}$$

- Based on this constant, we can write a valid probability density function as

$$\tilde{f}(x) = \frac{1}{k}\left(-x^2 + 2x\right) = \frac{3}{4}\left(-x^2 + 2x\right) \tag{21}$$

- Following back to Definition 3.3.2, suppose we want to define the probability density function such that $x \in \left[\frac{1}{2}, \frac{3}{4}\right]$.

$$\begin{aligned} \mathrm{P}\left[\frac{1}{2}, \frac{3}{4}\right] &= \frac{3}{4}\left(-\int_{\frac{1}{2}}^{\frac{3}{4}} x^2 dx + 2\int_{\frac{1}{2}} \frac{3}{4} x dx\right) \\ &= -\frac{1}{4}\left(\frac{27}{8} - \frac{1}{8}\right) + \frac{3}{4}\left(\frac{9}{4} - \frac{1}{4}\right) \\ &= \frac{22}{32} \end{aligned} \tag{22}$$

- The conditional probability over this range can then be written as

$$\tilde{f}(x) = \frac{\frac{3}{4}\left(-x^2 + 2x\right)}{\frac{22}{32}} \tag{23}$$

Some Common Univariate Distribution Functions

- Uniform Distributions

$$f\left(x\,|a,b\right)=\begin{cases}\frac{1}{b-a}\mathrm{ifa} \leq \mathrm{x} \leq \mathrm{b}\\ 0\ \ \mathrm{otherwise}\end{cases} \tag{24}$$

- Gamma Distribution

$$f\left(x\,|\alpha,\beta\right)=\begin{cases}\frac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{\alpha-1}e^{-{x}/{\beta}}\,\mathrm{given}\,0\ \ <\mathrm{x}<\infty,\alpha>0,\beta>0\\ \qquad 0\,\mathrm{otherwise}\end{cases} \tag{25}$$

- Normal Distribution

$$f\left(x\left|\mu,\sigma^2\right.\right)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]\text{ for all }-\infty<\mathrm{x}<\infty \tag{26}$$

- Beta Distribution

$$f\left(x\left|\alpha,\beta\right.\right)=\begin{cases}\frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}\text{ for }0<\mathrm{x}<1,\alpha>0,\beta>0\\ 0\text{ otherwise}\end{cases} \tag{27}$$

Readings

- Moss – 3.5 - 3.8
- Popper Chapter 2 – On the Problem of a Theory of Scientific Method