Wednesday, November 15 **Fall 2023**

Section 9: Transformations, Gamma/Beta, Conditional Expectation

Srihari Ganesh

Based on section note formatting template by Rachel Li and Ginnie Ma '23

1 Summary

1.1 Transformations

Definition 1 (Transformation)**.** A **transformation** of a random variable X is any (deterministic) function of that random variable $g(X)$.

Example 2 (Expo scaling)**.** Suppose $X \sim \text{Expo}(\lambda)$, and define $g(x) = \lambda x$. Then $\lambda X \sim \text{Expo}(1)$.

Example 3 (Uniform location-scale)**.** Suppose $U \sim \text{Unif}(0,1)$ and define $g(x) = (b-a)u + a$ for constants a, b. Then $g(U) = (b-a)U + a \sim \text{Unif}(a, b)$.

Result 4 (LOTUS)**.** Recall the **Law of the Unconscious Statistician (LOTUS)** to find the expectation of a transformed random variable $g(X)$ in terms of the PDF/PMF of X and function g.

$$E(X) = \begin{cases} \sum_{x \in S} g(x) P(X = x) & X \text{ discrete,} \\ \int_S g(x) f_X(x) dx & X \text{ continuous.} \end{cases}$$

Result 5 (CDF of transformed r.v.)**.** For *any* random variable X and *strictly increasing* function g, the CDF of $g(X)$ in terms of the CDF of X is

$$F_{g(X)}(y) = P(g(X) \le y) = P(X \le g^{-1}(y)) = F_X(g^{-1}(y)).$$

Result 6 (PMF of discrete transformed r.v.)**.** For *discrete* random variable X and *any* function g, the PMF of $g(X)$ in terms of the PMF of X is

$$P(g(X) = y) = \sum_{x : g(x) = y} P(X = x)$$

by LOTP.

Result 7 (Change of variables, PDF of continuous transformed r.v.)**.** For *continuous* random variable X and *differentiable, strictly monotone* function g, the PDF of $g(X)$ in terms of the PDF of X is

$$f_{g(X)}(y) = f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right|.$$

A strictly monotone function is either strictly increasing or strictly decreasing.

1.2 Gamma and Beta distributions

1.2.1 Distribution definitions and properties

Definition 8 (Beta distribution)**.** A random variable $B \sim \text{Beta}(a, b)$ that follows a **Beta distribution** has support $(0, 1)$ with the following PDF on the support:

$$f_B(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1},$$

where the integrating constant $\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$ uses Gamma functions as defined in Definition 9. Once you identify that the PDF of a random variable exclusively consists of a product of powers of x and $(1-x)$ with support $(0,1)$, you can uniquely identify it as a Beta distribution with parameters without having to wrangle the constants.

Definition 9 (Gamma function)**.** The **gamma function** Γ is defined for $a > 0$ by

$$\Gamma(a) = \int_0^\infty x^a e^{-x} \frac{1}{x} dx.$$

This does not have a general closed form and you will never have to solve it. Here are the handy properties:

- $\Gamma(a+1) = a\Gamma(a)$.
- If n is a positive integer, $\Gamma(n) = (n-1)!$.
- $\Gamma(1/2) = \sqrt{\pi}$.

Definition 10 (Gamma distribution)**.** A random variable that follows a **Gamma distribution** is denoted $G \sim \text{Gamma}(a, \lambda)$ with parameters $a, \lambda > 0$.

Result 11 (Gamma scaling)**.** If $G \sim \text{Gamma}(a, \lambda)$, then $\lambda G \sim \text{Gamma}(a, 1)$.

1.2.2 Distribution connections and stories

Result 12 (Beta-Uniform equivalence)**.** The $\text{Beta}(1,1)$ and $\text{Unif}(0,1)$ distributions are equivalent.

Result 13 (Beta-Uniform order statistics)**.** Suppose $U_1, \ldots, U_n \overset{i.i.d}{\sim} \text{Unif}(0,1)$. Then the k-th order statistic (the k-th smallest of the $\{U_i\}_{i=1,\ldots,n}$, as in Definition 19) follows the distribution $U_{(k)} \sim \text{Beta}(k, n-k+1)$. Note that Result 12 is consistent with this fact.

Result 14 (Beta-Binomial conjugacy)**.** Let $p \sim \text{Beta}(a, b)$ represent our prior belief about a probability. Suppose we observe $X|p \sim \text{Bin}(n, p)$, where n is constant but p is random. Then the posterior distribution of p (via Bayes' rule) is $p|X = k \sim \text{Bin}(a+k, b+n-k)$.

Result 15 (Bayes' billiards)**.** For integers k, n with $0 \le k \le n$,

$$\int_0^1 \binom{n}{k} x^k (1-x)^{n-k} dx = \frac{1}{n+1}.$$

Result 16 (Bank-Post Office story/Beta-Gamma connection)**.** Suppose $X \sim \text{Gamma}(a, \lambda)$ and $Y \sim \text{Gamma}(b, \lambda)$ with X, Y independent. Define $T = X + Y$ and $W = \frac{X}{X+Y}$. Then the following results hold:

1. $T \sim \text{Gamma}(a+b, \lambda)$,
2. $W \sim \text{Beta}(a, b)$, and
3. T and W are independent.

Result 17 (Expo-Gamma)**.** For $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Expo}(\lambda)$, $X_1 + \cdots X_n \sim \text{Gamma}(n, \lambda)$. Note that this means $\text{Expo}(\lambda)$ and $\text{Gamma}(1, \lambda)$ are the same distribution.

Result 18 (Poisson process-Gamma)**.** Suppose T_k is the k-th arrival time in a Poisson process with rate λ. Then by Result 17, $T_k \sim \text{Gamma}(k, \lambda)$.

1.3 Order statistics

Definition 19 (Order Statistics)**.** The **order statistics** of random variables $X_1, \ldots, X_n$ are the result of sorting them in increasing order, denoted $X_{(1)}, \ldots, X_{(n)}$. For example, $X_{(1)} = \min\{X_1, \ldots, X_n\}$ and $X_{(n)} = \min\{X_1, \ldots, X_n\}$. The **k-th order statistic**, notated $X_{(k)}$, is the k-th smallest value among $X_{(1)}, \ldots, X_{(n)}$.

☣ **20.** Order statistics are always going to be dependent by nature: If you know $X_{(1)}$, then you know $X_{(2)} \geq X_{(1)}$.

Result 21 (CDF of order statistic)**.** For $X_1, \ldots, X_n$ i.i.d. random variables with CDF F,

$$P(X_{(j)} \leq x) = \sum_{k=j}^{n} \binom{n}{k} F(x)^k (1 - F(x))^{n-k}.$$

Result 22 (PDF of order statistic)**.** For $X_1, \ldots, X_n$ i.i.d. *continuous* random variables with PDF f and CDF F,

$$f_{X_{(j)}}(x) = n \binom{n-1}{j-1} f(x) F(x)^{j-1} (1 - F(x))^{n-j}.$$

1.4 Conditional expectation

Definition 23 (Conditional expectation given event)**.** The **conditional expectation** of a random variable X given an event A is

$$E(X|A) = \begin{cases} \sum_x P(X = x|A) & X \text{ discrete,} \\ \int_{-\infty}^{\infty} x f(x|A) dx & X \text{ continuous.} \end{cases}$$

Definition 24 (Law of total expectation)**.** Suppose $A_1, A_2, \ldots$ are events partitioning the sample space. Then the **law of total expectation (LOTE)** states that

$$E(X) = \sum_j E(X|A_j) P(A_j).$$

2 Practice Problems

1. Suppose Shawn attempts a free throw, p is the probability that they make it. Let $p \sim \text{Unif}(0,1)$ be our prior distribution for that probability. Shawn attempts 5 free throws and makes 4 of them. Assuming that each free throw is i.i.d. (they have an equal and independent probability p of making each), what is the posterior distribution of p?

> **Solution**
> Let X be the number of free throws that Shawn makes out of 5. Then $X|p \sim \text{Bin}(5,p)$. Also recall that $p \sim \text{Beta}(1,1)$ (Result 12). So by the Beta-Binomial conjugacy (Result 14), $p|X = 4 \sim \text{Beta}(5,2)$.

2. Suppose $X \sim \text{Beta}(2,2)$. Find $E(X^3)$ by pattern-matching the LOTUS integral to a Beta PDF. Recall that for $x \in (0,1)$, a $\text{Beta}(a,b)$ r.v. has PDF

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}.$$

Fully simplify your answer (i.e., no Gamma functions left over, use the properties from Definition 9 if needed).

Solution

By LOTUS,

$$\begin{aligned} E(X^3) &= \int_0^1 x^3 \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)} x(1-x)dx \\ &= \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)} \int_0^1 x^4(1-x)dx \end{aligned}$$

Consider the PDF of a $\text{Beta}(5,2)$ random variable:

$$\begin{aligned} 1 &= \int_0^1 \frac{\Gamma(7)}{\Gamma(5)(\Gamma(2)} x^4(1-x)dx \\ \frac{\Gamma(5)\Gamma(2)}{\Gamma(7)} &= \int_0^1 x^4(1-x)dx. \end{aligned}$$

So plugging back in,

$$E(X^3) = \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)} \frac{\Gamma(5)\Gamma(2)}{\Gamma(7)}.$$

Recall that $\Gamma(n) = (n-1)!$ for positive integers n. So

$$\begin{aligned} E(X^3) &= \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)} \frac{\Gamma(5)\Gamma(2)}{\Gamma(7)} \\ &= \frac{\Gamma(4)\Gamma(5)}{\Gamma(2)\Gamma(7)} \\ &= \frac{3!4!}{1!6!} \\ &= \frac{3!}{6 \cdot 5} = \frac{1}{5}. \end{aligned}$$

3. *(inspired by HW 5.3)* Suppose George has $N \sim \text{Pois}(\lambda)$ children in his lifetime, and his i-th child has $G_i \sim \text{Pois}(\lambda)$ children themselves, and $G_1, \ldots, G_N, N$ are all independent. By pattern-matching to a Poisson expectation, find the expected number of grandchildren that George has.

Solution

To find this expectation, we wish we knew N. So using the law of total expectation,

$$E(G_1 + \cdots + G_N) = \sum_{n=0}^{\infty} E(G_1 + \cdots G_N | N = n) P(N = n).$$

See that $G_1 + \cdots + G_N | N = n \sim \text{Pois}(n\lambda)$. Also plug in $P(N = n)$ from the PMF of $N \sim \text{Pois}(\lambda)$.

$$E(G_1 + \cdots + G_N) = \sum_{n=0}^{\infty} (n\lambda) \frac{e^{-\lambda}\lambda^n}{n!}.$$

By definition,

$$\lambda = E(N) = \sum_{n=0}^{\infty} n \frac{e^{-\lambda}\lambda^n}{n!}.$$

So

$$E(G_1 + \cdots + G_N) = n\lambda.$$