Transformations, Beta, Gamma, Order Statistics
TFs: Justin Zhu

Continuous Transformations

Why do we need the Jacobian? We need the Jacobian to rescale our PDF so that it integrates to 1.

One Variable Transformations Let's say that we have a random variable X with PDF $f_X(x)$, but we are also interested in some function of X. We call this function $Y = g(X)$. Note that Y is a random variable as well. If g is differentiable and one-to-one (every value of X gets mapped to a unique value of Y), then the following is true:

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right| \text{ or } f_Y(y) \left| \frac{dy}{dx} \right| = f_X(x)$$

To find $f_Y(y)$ as a function of y, plug in $x = g^{-1}(y)$.

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

The derivative of the inverse transformation is referred to the **Jacobian**, denoted as J.

$$J = \frac{d}{dy} g^{-1}(y)$$

Two Variable Transformations Similarily, let's say we know the joint distribution of U and V but are also interested in the random vector (X, Y) found by $(X, Y) = g(U, V)$. If g is differentiable and one-to-one, then the following is true:

$$f_{X,Y}(x, y) = f_{U,V}(u, v) \left| \left| \frac{\delta(u, v)}{\delta(x, y)} \right| \right| = f_{U,V}(u, v) \left| \left| \begin{matrix} \frac{\delta u}{\delta x} & \frac{\delta u}{\delta y} \\ \frac{\delta v}{\delta x} & \frac{\delta v}{\delta y} \end{matrix} \right| \right| \text{ or } f_{X,Y}(x, y) \left| \left| \frac{\delta(x, y)}{\delta(u, v)} \right| \right| = f_{U,V}(u, v)$$

The outer $||$ signs around our matrix tells us to take the absolute value. The inner $||$ signs tells us to the matrix's determinant. Thus the two pairs of $||$ signs tell us to take the absolute value of the determinant matrix of partial derivatives. In a 2x2 matrix,

$$\left| \left| \begin{matrix} a & b \\ c & d \end{matrix} \right| \right| = |ad - bc|$$

The determinant of the matrix of partial derivatives is referred to the **Jacobian**, denoted as J.

$$\left| \begin{matrix} \frac{\delta u}{\delta x} & \frac{\delta u}{\delta y} \\ \frac{\delta v}{\delta x} & \frac{\delta v}{\delta y} \end{matrix} \right| = J$$

Gamma Function

The Letter Gamma - Γ is the (capital) roman letter Gamma. It is used in statistics for both the Gamma function and the Gamma Distribution.

Recursive Definition - The Gamma function is an extension of the factorial function to all real (and complex) numbers, with the argument shifted down by 1. When n is a positive integer,

$$\Gamma(n) = (n-1)!$$

For all values of n (except -1 and 0),
$$\Gamma(n+1) = n\Gamma(n)$$

Closed-form Definition - The Gamma function is defined as:

$$\Gamma(n) = \int_0^\infty t^{n-1}e^{-t}dt$$

Gamma Distribution (Continuous)

Let us say that X is distributed Gamma(a, λ). We know the following:

Story You sit waiting for shooting stars, and you know that the waiting time for a star is distributed Expo(λ). You want to see "a" shooting stars before you go home. X is the total waiting time for the ath shooting star.

Example You are at a bank, and there are 3 people ahead of you. The serving time for each person is distributed Exponentially with mean of 2 time units. The distribution of your waiting time until you begin service is Gamma$(3, \frac{1}{2})$

PDF The PDF of a Gamma is:

$$f(x) = \frac{1}{\Gamma(a)}(\lambda x)^a e^{-\lambda x}\frac{1}{x}, \quad x \in [0, \infty)$$

Properties and Representations

$$E(X) = \frac{a}{\lambda} \qquad Var(X) = \frac{a}{\lambda^2}$$

$$X \sim \text{Gamma}(a, \lambda), \ Y \sim \text{Gamma}(b, \lambda), \ X \perp\!\!\!\perp Y \ \longrightarrow \ X + Y \sim \text{Gamma}(a+b, \lambda), \ \frac{X}{X+Y} \perp\!\!\!\perp X + Y$$

$$X \sim \text{Gamma}(a, \lambda) \ \longrightarrow \ X = X_1 + X_2 + ...X_a \text{ for } X_i \text{ i.i.d. Expo}(\lambda)$$

$$\text{Gamma}(1, \lambda) \ \sim \ \text{Expo}(\lambda)$$

Beta Distribution (Continuous)

Let us say that X is distributed $\text{Beta}(a, b)$. We know the following:

Story (Bank/Post-Office) Let's say that your waiting time at the bank is distributed $X \sim \text{Gamma}(a, \lambda)$ and that your total waiting time at the post-office is distributed $Y \sim \text{Gamma}(b, \lambda)$. You visit both of them while doing errands. Your total waiting time at both is $X + Y \sim \text{Gamma}(a + b, \lambda)$ and the fraction of your time that you spend waiting at the Bank is $\frac{X}{X+Y} \sim \text{Beta}(a, b)$. The fraction is not dependent on λ, and $\frac{X}{X+Y} \perp\!\!\!\perp X + Y$.

Example You are tasked with finding Jules, Vernes, and Nemo in a friendly game of hide and seek. You look for them in order, and the time it takes to find one of them is distributed Exponentially with a mean of $\frac{1}{3}$ of a time unit. The time it takes to find both Jules and Vernes is $\text{Expo}(3) + \text{Expo}(3) \sim \text{Gamma}(2, 3)$. The time it takes to find Nemo is $\text{Expo}(3) \sim \text{Gamma}(1, 3)$. Thus the proportion of the total hide-and-seek time that you spend finding Nemo is distributed $\text{Beta}(1, 2)$ and is independent from the total time that you've spent playing the game.

PDF The PDF of a Beta is:

$$f(x) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1 - x)^{b-1}, \quad x \in (0, 1)$$

Properties and Representations

$$E(X) = \frac{a}{a + b}$$

$$X \sim \text{Gamma}(a, \lambda), \quad Y \sim \text{Gamma}(b, \lambda), \quad X \perp\!\!\!\perp Y \quad \longrightarrow \quad \frac{X}{X + Y} \sim \text{Beta}(a, b), \quad \frac{X}{X + Y} \perp\!\!\!\perp X + Y$$

$$\text{Beta}(1, 1) \sim \text{Unif}(0, 1)$$

Notable Uses of the Beta Distribution

...as the Order Statistics of the Uniform - The smallest of three Uniforms is distributed $U_{(1)} \sim \text{Beta}(1, 3)$. The middle of three Uniforms is distributed $U_{(2)} \sim \text{Beta}(2, 2)$, and the largest $U_{(3)} \sim \text{Beta}(3, 1)$. The distribution of the the jth order statistic of n i.i.d Uniforms is:

$$U_{(j)} \sim \text{Beta}(j, n - j + 1)$$

$$f_{U_{(j)}}(u) = \frac{n!}{(j - 1)!(n - j)!} t^{j-1}(1 - t)^{n-j}$$

...as the Conjugate Prior of the Binomial - A prior is the distribution of a parameter before you observe any data ($f(x)$). A posterior is the distribution of a parameter after you observe data y ($f(x|y)$). Beta is the *conjugate* prior of the Binomial because if you have a Beta-distributed prior on p (the parameter of the Binomial), then the posterior distribution on p given observed data is also Beta-distributed. This means, that in a two-level model:

$$X|p \sim \text{Bin}(n, p)$$
$$p \sim \text{Beta}(a, b)$$

Then after observing the value $X = x$, we get a posterior distribution $p|(X = x) \sim \text{Beta}(a + x, b + n - x)$

Special Cases of Beta and Gamma

$$\text{Gamma}(1, \lambda) \sim \text{Expo}(\lambda) \qquad \text{Beta}(1, 1) \sim \text{Unif}(0, 1)$$

Bank and Post Office Result

Let us say that we have $X \sim \text{Gamma}(a, \lambda)$ and $Y \sim \text{Gamma}(b, \lambda)$, and that $X \perp\!\!\!\perp Y$. By Bank-Post Office result, we have that:

$$X + Y \sim \text{Gamma}(a + b, \lambda)$$
$$\frac{X}{X + Y} \sim \text{Beta}(a, b)$$
$$X + Y \perp\!\!\!\perp \frac{X}{X + Y}$$

Order Statistics

Definition - Let's say you have n i.i.d. random variables $X_1, X_2, X_3, \ldots X_n$. If you arrange them from smallest to largest, the ith element in that list is the ith order statistic, denoted $X_{(i)}$. $X_{(1)}$ is the smallest out of the set of random variables, and $X_{(n)}$ is the largest.

Properties - The order statistics are dependent random variables. The smallest value in a set of random variables will always vary and itself has a distribution. For any value of $X_{(i)}$, $X_{(i+1)} \geq X_{(j)}$.

Distribution - Taking n i.i.d. random variables $X_1, X_2, X_3, \ldots X_n$ with CDF $F(x)$ and PDF $f(x)$, the CDF and PDF of $X_{(i)}$ are as follows:

$$F_{X_{(i)}}(x) = P(X_{(i)} \leq x) = \sum_{k=i}^{n} \binom{n}{k} F(x)^k (1 - F(x))^{n-k}$$
$$f_{X_{(i)}}(x) = n \binom{n-1}{i-1} F(x)^{i-1} (1 - F(X))^{n-i} f(x)$$

Universality of the Uniform - We can also express the distribution of the order statistics of n i.i.d. random variables $X_1, X_2, X_3, \ldots X_n$ in terms of the order statistics of n uniforms. We have that

$$F(X_{(j)}) \sim U_{(j)}$$

Practice Problems

Example 1. Transformations of the Uniform.
Let $U \sim \text{Unif}(0, 1)$. Find the PDF's of the following transformations:

(a) U^2

(b) $\sqrt{U}$

(c) $-\log(U)$

Example 2. Transformations of Exponentials.
Let $U, V \sim \text{Expo}(1)$ be two independent random variables. Define $X = U + V$ and $Y = \frac{U}{U+V}$.

(a) Find the joint PDF of X and Y

(b) Find the marginal distributions of X and Y

Example 3. Box-Muller Transformation.

Suppose we are told that $U \sim \text{Unif}(0,1)$ and $V \sim \text{Expo}\left(\frac{1}{2}\right)$ independently. Find the density function and thus the joint distribution of

$$X = \sqrt{V}\sin(2\pi U)$$
$$Y = \sqrt{V}\cos(2\pi U)$$

Example 4. More Beta Properties (BH 8.29).

Let $B \sim \text{Beta}(\alpha, \beta)$. Find the distribution of $1 - B$ in two ways by

(a) Using a change of variables

(b) Using a story proof related to the Gamma distribution.

Also explain why the result makes sense in terms of Beta being the conjugate prior for the Binomial.

Example 5. Second-best.
Let $U_1, \ldots, U_n$ be i.i.d. Unif$(0,1)$. Find the unconditional distribution of $U_{(n-1)}$, and the conditional distribution of $U_{(n-1)}$ given $U_{(n)} = c$.

Example 6. Order Statistics.
Let $X \sim \text{Bin}(n, p)$ and $B \sim \text{Beta}(j, n - j + 1)$, where n is a positive integer and j is a positive integer with $j \leq n$. Show using a story about order statistics that

$$P(X \geq j) = P(B \leq p)$$

This shows that the CDF of the continuous r.v. B is closely related to the CDF of the discrete r.v. X, and is another connection between the Beta and Binomial.