Stat 134: Change of Variable/Operations Review Session

December 5th, 2018

Conceptual Review

a. Write out the five steps for computing a change of variable density.

For computing the density of Y = f(X):

Step 1: Find the Range of Y

Step 2: Rewrite X in terms of Y

Step 3: Find dY/dX

Step 4: Plug density of X in change-of-variable formula

Step 5: Rewrite X in terms of Y

b. What is the formula for Z = X + Y, where X, Y are continuous random variables?

$$\int_{-\infty}^{\infty} f_{xy}(x, z-x)dx$$

c. What is the formula for Z = X/Y where X, Y are non-negative continuous random variables?

$$\int_{0}^{\infty} f_{xy}(zy, y)ydy$$

Problem 1: Properties of Exponentials

Suppose X is an Exponential(λ) and Y is a Gamma(n, λ). Let W = X/Y.

a. Find the distribution of cX

For computing the density of Z = cX:

Step 1: Range of Z: [0, ∞)

Step 2: Rewrite X in terms of Z: X = Z/c

Step 3: dZ/dX = c

Step 4: $f_z(z) = f_x(x) / |dZ/dX|$
$f_z(z) = \lambda e^{-\lambda x}$ / c

Step 5: Rewrite X in terms of Z: $f_z(z) = (\lambda/c)e^{-\lambda z/c}$

Z is Exp(λ/c)

b. Find the distribution of cY

Y = W_1 + W_2 + .. W_n where W_i is Exp(λ)

cY = cW_1 + cW_2 + .. cW_n

Using (a), cY is the sum of n Exp(λ/c)

Therefore, cY is a Gamma(n, λ/c)

c. Use the above parts to find the CDF of W

If $n \in \mathbb{N}$ and X, Y are independent, $F_W(w) = \mathbb{P}[W \leq w] = \mathbb{P}[X \leq wY]$. From previous parts wY follows Gamma(n,$\frac{\lambda}{w}$). In this case, the probability $\mathbb{P}[X \leq wY]$ has a nice interpretation.

Consider X as the waiting time till the first red car (rate = λ) arrives, and wY as the waiting time till the nth green car (rate = $\frac{\lambda}{w}$) arrives. The probability mentioned above is the probability that the first red car arrives before the nth green car, which is equal to $1 - \mathbb{P}[\text{first red after nth green}]$. Apply competing exponential n times, one gets $1 - (\frac{\lambda}{\lambda + w\lambda})^n$.

If the condition $n \in \mathbb{N}$ is not satisfied, the same expression can be obtained by taking double integral (given X, Y are independent).

$$
\begin{aligned}
\mathbb{P}[X \leq wY] &= \int_0^\infty \int_0^{wy} f_X(x) fY(y) dxdy \\
&= \int_0^\infty fY(y) \int_0^{wy} f_X(x) dxdy \\
&= \int_0^\infty fY(y) F_x(wy) dy \\
&= \int_0^\infty fY(y)(1 - e^{-\lambda wy}) dy \\
&= 1 - \int_0^\infty \frac{1}{\Gamma(n)} y^{n-1} \lambda^n e^{-(\lambda + w\lambda)y} dy \\
&= 1 - \frac{\lambda^n}{(\lambda + w\lambda)^n} \int_0^\infty \frac{1}{\Gamma(n)} y^{n-1} (\lambda + w\lambda)^n e^{-(\lambda + w\lambda)y} dy \\
&= 1 - \frac{\lambda^n}{(\lambda + w\lambda)^n}.
\end{aligned}
$$

The last equality holds since the integrand is the probability density function of Gamma(n,$\lambda + w\lambda$), and integrating over its support gives value 1.

Problem 2

Let X be the minimum of 6 independent Uniform(0,1) distributions and Y be the maximum. Find the distribution of $Z = X/Y$ and $W = Y - X$

We need to start by computing the joint density of X, Y

f_{xy}(x, y) = 6 * 5 * $(y-x)^4$

f_{xy}(x, y) = 30$(y-x)^4$

Now, we can use the ratio convolution formula to compute the density of Z

$$\begin{aligned} f_z(z) &= \int_0^1 f_{xy}(zy, y) y dy \\ &= \int_0^1 30(y - yz)^4 y dy \\ &= \int_0^1 30(1-z)^4 y^5 dy. \end{aligned}$$

Solving this integral gives us f_Z(z) = 5(1-z)^4

We can use the convolution formula to compute the density of W

$$\begin{aligned} f_w(w) &= \int_w^1 f_{xy}(y-w, y) dy \\ &= \int_w^1 30(y-(y-w))^4 dy \\ &= \int_w^1 30w^4 dy. \end{aligned}$$

Solving this integral gives us $\mathrm{f}_w(w) = 30\mathrm{w}^4(1-w)$.

Problem 3

Let $U_1, U_2, ...U_n$ be independent Uniform(0,1) random variables.

a. Find the distribution of $-log(U_1U_2...U_n)$

b. Use the above to find the probability $-log(U_1U_2...U_{10}) > 8$

a.
Notice that $-\log(U) \sim$ Exp(1), where U is Unif(0,1).
By property of logarithm, we have $-\log(U_1U_2\cdots U_n) = -\log(U_1) - \log(U_2) - \cdots - \log(U_n) \sim E_1 + \cdots + E_n \sim$ Gamma(n,1), where $E_1, \cdots, E_n$ are i.i.d exponential random variables with rate 1.
Therefore, the distribution is Gamma(n,1).
b.
Since $-\log(U_1\cdots U_{10})$ follows a Gamma(10,1) distribution,

$$\begin{aligned}\mathbb{P}[-\log(U_1\cdots U_{10}) > 8] &= \mathbb{P}[]N_{(0,8)} \leq 9] \\ &= \sum_{k=0}^{9} e^{-8}\frac{8^k}{k!}.\end{aligned}$$