Statistical and Mathematical Methods DS-GA 1002
, 2019

Recitation Session 3 Solutions

Problem

1. *Calculus.*
   a. Calculate the derivative of $y = (\log(x))^{\log(x)}$
   b. Calculate $\int_0^\infty e^{\frac{-x^2}{2}} dx$

   SOLUTION (10 points)

   a. Calculate the derivative of $y = (\log(x))^{\log(x)}$
   Let $u = log(x)$. Then,

$$
\begin{aligned}
y &= u^u && (1)\\
ln(y) &= ln(u^u) && (2)\\
\frac{1}{y}\frac{dy}{du} &= ln(u) + u\frac{1}{u} && (3)\\
\frac{dy}{du} &= y \cdot (ln(u) + u\frac{1}{u}) && (4)\\
\frac{dy}{du} &= log(x)^{log(x)} \cdot (ln(log(x)) + 1) && (5)\\
\frac{du}{dx} &= \frac{1}{x} && (6)\\
\frac{dy}{dx} &= \frac{dy}{du} \cdot \frac{du}{dx} && (7)\\
\frac{dy}{dx} &= log(x)^{log(x)} \cdot (ln(log(x)) + 1) \cdot \frac{1}{x} && (8)
\end{aligned}
$$

   b. Calculate $\int_0^\infty e^{\frac{-x^2}{2}} dx$

   Let $I = \int_{-\infty}^\infty e^{\frac{-x^2}{2}} dx$
   Let $x^2 + y^2 = r^2 \Rightarrow dxdy = rd\theta dr$
   Let $u = \frac{r^2}{2} \Rightarrow du = \frac{2r}{2}dr = rdr$

$$
\begin{aligned}
I^2 &= (\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx)^2 && (9)\\
&= \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx \int_{-\infty}^{\infty} e^{\frac{-y^2}{2}} dy && (10)\\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\frac{-x^2-y^2}{2}} dxdy && (11)\\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\frac{x^2+y^2}{-2}} dxdy && (12)\\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\frac{r^2}{-2}} r d\theta dr && (13)\\
&= \int_{0}^{2\pi} \int_{0}^{\infty} e^{\frac{r^2}{-2}} r dr d\theta && (14)\\
&= 2\pi \int_{r=0}^{\infty} e^{\frac{r^2}{-2}} r dr && (15)\\
&= 2\pi \int_{u=0}^{\infty} e^{-u} du && (16)\\
&= 2\pi(-e^{-\infty} - (-e^0)) && (17)\\
&= 2\pi && (18)\\
I &= \sqrt{2\pi} && (19)\\
\int_{0}^{\infty} e^{\frac{-x^2}{2}} dx &= \frac{1}{2}(\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx) && (20)\\
&= \frac{I}{2} && (21)\\
&= \frac{\sqrt{2\pi}}{2} && (22)
\end{aligned}
$$

2. *Change of variables.*

   a. Let X be uniformly distributed over $[0, 1]$, and $Y = -\lambda^{-1} \log(X)$ where λ is positive. Find the distribution of Y

   b. Let $X \sim N(0, 1)$ and $Y = e^X$. Find the pdf of Y

   SOLUTION (10 points)

   a. Main idea for such problems is that probabilities measure sizes of sets, hence if the set is unchanged then probabilities are also unchanged. In other words $\mathrm{P}(X \in E) = \mathrm{P}(f(X) \in f(E))$ note the slight abuse of notation here, what $X \in E$ means is $\{\omega \in \Omega | X(\omega) \in E\}$. Again with a slight abuse of notation we can write $\mathrm{P}(X \in dx) = \mathrm{P}(Y \in dy)$ which gives us: $f_Y(y) = f_X(x)/|dy/dx|$. Absolute values are there because P measures a quantity that is always positive.

$$\frac{dy}{dx} = \frac{1}{\lambda x} \tag{23}$$

$$f_Y(y) = \frac{1}{1/\lambda x} = \lambda x \tag{24}$$

$$= \lambda \exp\left(-\lambda y\right) \tag{25}$$

$$\tag{26}$$

It turns out that Y is an exponential distribution with parameter λ. This is a standard way to simulate exponential distribution using uniform distribution.

b.

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$$

$$\mathbb{P}(X \le x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\ (-\ \frac{t^2}{2})\mathrm{d}t$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-\frac{t^2}{2})\mathrm{d}t$$

$$\mathbb{P}(Y \le y) = \mathbb{P}(\exp(X) \le y)$$

$$= \mathbb{P}(X \le \ln y)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\ln y} \exp(-\frac{t^2}{2})\mathrm{d}t,\ 0 < y < \infty$$

Denote $q(t) = \exp(-\ \frac{t^2}{2}); Q(t) = \int q(t)\mathrm{d}t$

$$\frac{\mathrm{d}}{\mathrm{d}y}\mathbb{P}(Y \le y) = \frac{\mathrm{d}}{\mathrm{d}y}\frac{1}{\sqrt{2\pi}}(Q(\ln y) - Q$$

$$= \frac{1}{\sqrt{2\pi}}\frac{\mathrm{d}}{\mathrm{d}y}Q(\ln y)$$

$= \frac{1}{\sqrt{2\pi}}Q'(\ln y)\frac{\mathrm{d}}{\mathrm{d}y}\ln y$ by chain rule

$$= \frac{1}{\sqrt{2\pi}}q(\ln y)\frac{1}{y}$$

$$= \frac{1}{y\sqrt{2\pi}}\exp(-\frac{(\ln y)^2}{2})$$

$$f_Y(y) = \frac{1}{y\sqrt{2\pi}}\exp(-\frac{(\ln y)^2}{2})\ ,\ 0 < y < \infty$$

3. *Inverse transform: discrete RV.* Figure 1 shows the cdf of a Binomial(2, 0.5) random variable.

   a. How would you obtain this cdf as a function $g(U)$ from the uniform $(0, 1)$ variable U?

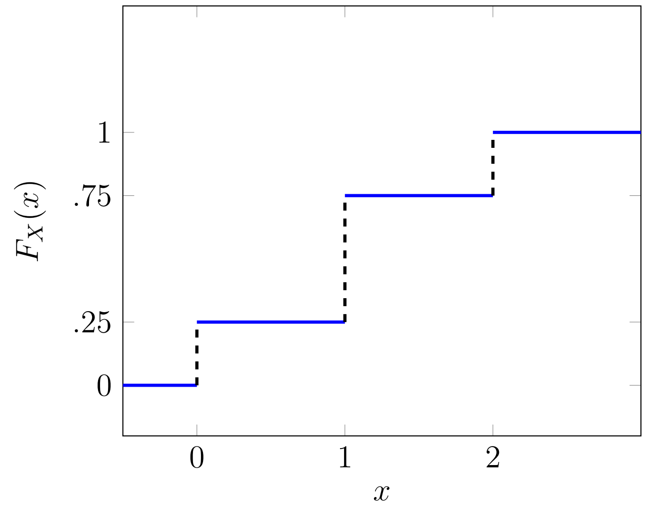

Figure 1: cdf of a binomial (2, 0.5) random variable.

b. How would you obtain $Y \sim$ Binomial(10, 0.5) from (a)?

c. How would you simulate it using fair coin flips?

Solution

a. The rule for getting from the uniform $(0,1)$ variable U to the binomial (2,0.5) variable $g(U)$ is

$$g(u) = \begin{cases} 0 & \text{if } 0 \leq u \leq 0.25 \\ 1 & \text{if } 0.25 < u \leq 0.75 \\ 2 & \text{if } 0.75 < u \leq 1 \end{cases}$$

This is because by construction the intervals on which g takes the values 0,1,2 have lengths $0.25, 0.5$, and 0.25 respectively, as required by the binomial $(2, 0.5)$ distribution.

b.

$$\{X_i\}_{i=1^{iid}}^{5} \sim \text{Binomial}(2, 0.5)$$

$$Y = \sum_{i=1}^{5} X_i$$

We can obtain a realization of Y by re-doing (b) 5 times and sum the results together.

$$\Rightarrow g(u_1) + g(u_2) + g(u_3) + g(u_4) + g(u_5) = y_1.$$

c. To simulate it using fair coin flips, use

$$g(u) = \begin{cases} 0 & \text{if HH} \\ 1 & \text{if HT or TH} \\ 2 & \text{if TT} \end{cases}$$

4. *Laplace Distribution.* A continuous random variable X is said to have a Laplace distribution with parameter λ if its pdf is given by:

$$f(x) = A\exp(-\lambda|x|), \quad -\infty < x < \infty$$

for some constant A.

a. Can the parameter λ be negative or zero

b. Compute the constant A in terms of λ

c. Compute the cdf of X

d. For $s, t > 0$, compute $p(X > s + t|X > s)$

e. Compute the pdf of $Y = |X|$

SOLUTION (10 points)

a. No. We know that for a valid pdf,

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

If we plug in for $f(x)$, we get

$$\int_{-\infty}^{\infty} A \ \exp(-\lambda|x|)dx = 2\int_{0}^{\infty} A \ \exp(-\lambda x)dx = 2[\frac{-A}{\lambda}\exp(-\lambda x)]_0^\infty$$

If $\lambda \leq 0$, this evaluates to infinity, and then there is no way to choose A so that the area under the pdf is 1. Therefore, it must be that $\lambda > 0$.

b. Following part (a), we know

$$\int_{-\infty}^{\infty} A \ \exp(-\lambda|x|)dx = 2\int_{0}^{\infty} A \ \exp(-\lambda x)dx = 1 \ .$$

Also,

$$2\int_0^\infty A\exp(-\lambda x)dx = 2[\frac{-A}{\lambda}\exp(-\lambda x)]_{x=0}^\infty$$

$$= 0 - 2 \times (-\frac{A}{\lambda})$$

$$= 2\frac{A}{\lambda}.$$

Therefore, we must have $2A/\lambda = 1$, or equivalently, $A = \lambda/2$.

c.

$$\mathrm{P}(X \leq x) = \int_{-\infty}^{x} f(u)du$$

$$= \int_{-\infty}^{x} \frac{\lambda}{2}\exp(-\lambda|u|)du$$

When $x \le 0$,

$$\begin{aligned} P(X \le x) &= \int_{-\infty}^{x} \frac{\lambda}{2} \exp(-\lambda(-u))du \\ &= \frac{1}{2}[\exp(\lambda u)]_{-\infty}^{x} \\ &= \frac{1}{2}\exp(\lambda x) \ . \end{aligned}$$

When $x > 0$,

$$\begin{aligned} P(X \le x) &= \int_{0}^{x} \frac{\lambda}{2} \exp(-\lambda u)du + \frac{1}{2} \\ &= \frac{1}{2}[-\exp(-\lambda u)]_{0}^{x} + \frac{1}{2} \\ &= \frac{1}{2}(-\exp(-\lambda x) + 1) + \frac{1}{2} \\ &= 1 - \frac{1}{2}\exp(-\lambda x) \ . \end{aligned}$$

d. We begin with the definition of conditional probability:

$$P(X \ge s + t | X \ge s) = \frac{P(X \ge s + t \cap X \ge s)}{P(X \ge s)} \ .$$

Since $t > 0$, if $X \ge s + t$ then $X \ge t$, so we have

$$P(X \ge s + t \cap X \ge s) = P(X \ge s + t) \ .$$

Also, since s and t are both positive, $P(X \ge s + t) = \frac{1}{2}\exp(-\lambda(s + t))$ and $P(X \ge s) = \frac{1}{2}\exp(-\lambda s)$. Therefore,

$$P(X \ge s + t | X \ge s) = \frac{P(X \ge s + t)}{P(X \ge s)} = \exp(-\lambda t) \ .$$

e. Note that for $y \ge 0$,

$$P(Y \le y) = 1 - P(Y > y) = 1 - (P(X > y) + P(X < -y)) \ .$$

Thus,

$$P(Y \le y) = 1 - (\frac{1}{2}\exp(-\lambda y) + \frac{1}{2}\exp(-\lambda y)) = 1 - \exp(-\lambda y) \ .$$

This is the cumulative distribution function for Y. Finally, we differentiate $P(Y \le y)$ to get the pdf of Y, noting that

$$\frac{d}{dy}(1 - \exp(-\lambda y)) = \lambda \exp(-\lambda y) \ .$$

Consequently, the pdf for Y is given by

$$f(y) = \begin{cases} \lambda \exp(-\lambda y) & \text{if } y \ge 0, \\ 0 & \text{if } y < 0. \end{cases}$$