Tutorial 3

October 8, 2020

Question 1

Let N be a Poisson random variable with parameter b, and consider a sequence of N independent Bernoulli trials, each with probability p for success.

Let X be the total number of successes. Find the distribution of X.

$$
\begin{aligned}
\mathbf{P}(X=k) &= \sum_{n=0}^{\infty} \mathbf{P}(X=k \mid N=n) \cdot \mathbf{P}(N=n) && \text{(Law of total probability)} \\
&= \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} e^{-b} \frac{b^n}{n!} && \text{(Must have } n \geq k\text{)} \\
&= \sum_{n=k}^{\infty} \frac{n!\, p^k (1-p)^{n-k} e^{-b} b^n}{(n-k)!\, k!\, n!} && \text{(Expand binomial coefficient)} \\
&= e^{-b} \frac{p^k}{k!} \sum_{n=k}^{\infty} \frac{b^n (1-p)^{n-k}}{(n-k)!} && \text{(}n!\text{ cancels)} \\
&= e^{-b} \frac{b^k p^k}{k!} \sum_{n=k}^{\infty} \frac{b^{n-k} (1-p)^{n-k}}{(n-k)!} && \text{(Factor } b^k \text{ out of the sum)} \\
&= e^{-b} \frac{b^k p^k}{k!} \sum_{n-k=0}^{\infty} \frac{b^{n-k} (1-p)^{n-k}}{(n-k)!} && \text{(Sum index now starts at zero)} \\
&= e^{-b} \frac{b^k p^k}{k!} e^{b(1-p)} && \text{(Power series definition of } e^x\text{)} \\
&= e^{-bp} \frac{(bp)^k}{k!}
\end{aligned}
$$

We recognize this as the PDF for Poisson(bp). Therefore $X \sim \text{Poisson}(bp)$.

Question 2

A random variable X has PDF:

$$
f(x) = \begin{cases} cxe^{-x} & x > 0 \\ 0 & \text{otherwise} \end{cases}
$$

(a) Find the value of c which makes $f(x)$ a valid PDF.

For $f(x)$ to be a valid PDF, the integral of $f(x)$ over its support must be equal to 1.

$$\begin{aligned}\int_0^\infty cxe^{-x}\,dx &= c\int_0^\infty xe^{-x}\,dx \\ &= c\left(\lim_{t\to\infty} -xe^{-x} - e^{-x}\Big|_{x=0}^{x=t}\right) \\ &= c\left((0-0)-(0-1)\right) \\ &= c\end{aligned}$$

Therefore we have that c must be equal to 1. Our updated PDF is:

$$f(x) = \begin{cases} xe^{-x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

(b) Find the CDF of X.

We can obtain the CDF by integrating the PDF from negative infinity up to x, with respect to a dummy variable, t. Denote the CDF by $F(x)$.

$$\begin{aligned}F(x) &= \int_{-\infty}^x f(t)\,dt \\ &= \int_0^x te^{-t}\,dt \\ &= -te^{-t} - e^{-t}\Big|_{t=0}^{t=x} \\ &= (-xe^{-x} - e^{-x}) - (0-1) \\ &= 1 - e^{-x}(1+x)\end{aligned}$$

The full CDF is:

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-x}(1+x) & x > 0 \end{cases}$$

Question 3

Consider a random variable, X, with a triangular distribution:

$$f(x) = \begin{cases} x & 0 < x < 1 \\ 2 - x & 1 \leq x < 2 \\ 0 & \text{otherwise} \end{cases}$$

Find the mean and variance.

$$\mathbf{E}(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

$$= \int_0^1 x^2\, dx + \int_1^2 2x - x^2\, dx$$

$$= \frac{1}{3}x^3 \Big|_{x=0}^{x=1} + \left(x^2 - \frac{1}{3}x^3\right)\Big|_{x=1}^{x=2}$$

$$= \frac{1}{3} + \left(\left(4 - \frac{8}{3}\right) - \left(1 - \frac{1}{3}\right)\right)$$

$$= 1$$

To find the variance we could compute $\mathbf{E}\left((X-\mu)^2\right)$, but this integral can get messy. It is easier to use the relationship that

$$\mathbf{Var}(X) = \mathbf{E}\left(X^2\right) - \left(\mathbf{E}(X)\right)^2$$

and compute $\mathbf{E}\left(X^2\right)$ since we already computed $\mathbf{E}(X)$ in the previous step.

$$\mathbf{E}\left(X^2\right) = \int_{-\infty}^{\infty} x^2\, f(x)\, dx$$

$$= \int_0^1 x^3\, dx + \int_1^2 2x^2 - x^3\, dx$$

$$= \frac{1}{4}x^4 \Big|_{x=0}^{x=1} + \left(\frac{2}{3}x^3 - \frac{1}{4}x^4\right)\Big|_{x=1}^{x=2}$$

$$= \frac{1}{4} + \left(\left(\frac{16}{3} - \frac{16}{4}\right) - \left(\frac{2}{3} - \frac{1}{4}\right)\right)$$

$$= \frac{7}{6}$$

$$\mathbf{Var}(X) = \mathbf{E}\left(X^2\right) - \mathbf{E}(X)^2 = \frac{7}{6} - (1)^2 = \frac{1}{6}$$

Question 4

A random variable, X, has PDF:

$$f(x) = \begin{cases} x/2 & 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

Find $\mathbf{P}(X > 1.5 \mid X > 1)$.

First find the CDF.

$$F(x) = \int_0^x \frac{t}{2}\,dt = \left.\frac{t^2}{4}\right|_{t=0}^{t=x} = \frac{x^2}{4}$$

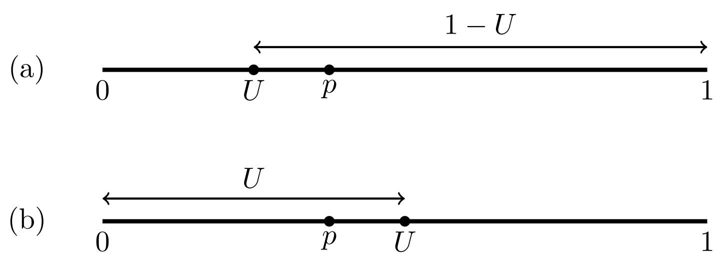
$$F(x) = \begin{cases} 0 & x < 0 \\ x^2/4 & 0 \le x \le 2 \\ 1 & x > 2 \end{cases}$$

$$\begin{aligned} \mathbf{P}(X > 1.5 \mid X > 1) &= \frac{\mathbf{P}(X > 1.5 \cap X > 1)}{\mathbf{P}(X > 1)} \\ &= \frac{\mathbf{P}(X > 1.5)}{\mathbf{P}(X > 1)} \\ &= \frac{1 - \mathbf{P}(X \le 1.5)}{1 - \mathbf{P}(X \le 1)} \\ &= \frac{1 - 9/16}{1 - 1/4} \\ &= \frac{7}{12} \approx 0.5833 \end{aligned}$$

Alternatively, we could have computed $\mathbf{P}(X > 1.5)$ as $\int_{1.5}^{2} f(x)\,dx$, and similarly with $\mathbf{P}(X > 1)$.

Question 5

A stick of length 1 is split at a point U that is uniformly distributed over (0,1). Determine the expected length of the piece that contains the point p, $0 \le p \le 1$.



There are two cases to consider: (a) $U < p$; (b) $U > p$.

Let $L_p(U)$ denote the length of the piece containing point p. Then:

$$L_p(U) = \begin{cases} 1 - U & U < p \\ U & U > p \end{cases}$$

Quick facts:

(i) The PDF of a Unif$(0, 1)$ random variable is 1

(ii) If U has density denoted by $p(u)$, then

$$\mathbf{E}\left(g(U)\right) = \int_{-\infty}^{\infty} g(u)p(u)\,du$$

If $U \sim \text{Unif}(0,1)$, then $p(u) = 1$ and the expectation above reduces to

$$\mathbf{E}\left(g(U)\right) = \int_{-\infty}^{\infty} g(u)\,du$$

Using the above facts, the expected length of the piece containing point p can be computed as:

$$\begin{aligned}
\mathbf{E}\left(L_p(U)\right) &= \int_{-\infty}^{\infty} L_p(u)\,du \\
&= \int_{0}^{p} (1-u)\,du + \int_{p}^{1} u\,du \\
&= \left(u - \frac{u^2}{2}\right)\Bigg|_{u=0}^{u=p} + \frac{1}{2}u^2\Bigg|_{u=p}^{u=1} \\
&= \left(p - \frac{p^2}{2}\right) + \left(\frac{1}{2} - \frac{p^2}{2}\right) \\
&= \frac{1}{2} + p(1-p)
\end{aligned}$$

It can be seen that the expected length of the piece containing point p reaches its maximum of $3/4$ when $p = 1/2$, since $p(1-p)$ is at its maximum here.