ECS 315: Probability and Random Processes **2016/1**

HW Solution 6 — Due: Not Due

Lecturer: Prapun Suksompong, Ph.D.

Problem 1. The thickness of the wood paneling (in inches) that a customer orders is a random variable with the following cdf:

$$F_X(x) = \begin{cases} 0, & x < \frac{1}{8} \\ 0.2, & \frac{1}{8} \le x < \frac{1}{4} \\ 0.9, & \frac{1}{4} \le x < \frac{3}{8} \\ 1 & x \ge \frac{3}{8} \end{cases}$$

Determine the following probabilities:

(a) $P[X \le 1/18]$

(b) $P[X \le 1/4]$

(c) $P[X \le 5/16]$

(d) $P[X > 1/4]$

(e) $P[X \le 1/2]$

[Montgomery and Runger, 2010, Q3-42]

Solution:

(a) $P[X \le 1/18] = F_X(1/18) = 0$ because $\frac{1}{18} < \frac{1}{8}$.

(b) $P[X \le 1/4] = F_X(1/4) = 0.9$

(c) $P[X \le 5/16] = F_X(5/16) = 0.9$ because $\frac{1}{4} < \frac{5}{16} < \frac{3}{8}$.

(d) $P[X > 1/4] = 1 - P[X \le 1/4] = 1 - F_X(1/4) = 1 - 0.9 = 0.1$.

(e) $P[X \le 1/2] = F_X(1/2) = 1$ because $\frac{1}{2} > \frac{3}{8}$.

Alternatively, we can also derive the pmf first and then calculate the probabilities.

Problem 2. [M2011/1] The cdf of a random variable X is plotted in Figure 6.1.

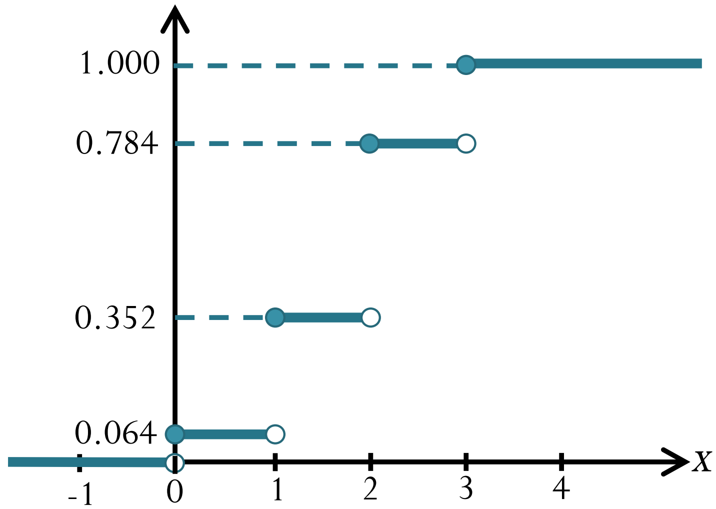

Figure 6.1: CDF of X for Problem 2

(a) Find the pmf $p_X(x)$.

(b) Find the family to which X belongs? (Uniform, Bernoulli, Binomial, Geometric, Poisson, etc.)

Solution:

(a) For discrete random variable, $P[X = x]$ is the jump size at x on the cdf plot. In this problem, there are four jumps at 0, 1, 2, 3.

- $P[X = 0] =$ the jump size at $0 = 0.064 = \frac{64}{1000} = (4/10)^3 = (2/5)^3$.
- $P[X = 1] =$ the jump size at $1 = 0.352 - 0.064 = 0.288$.
- $P[X = 2] =$ the jump size at $2 = 0.784 - 0.352 = 0.432$.
- $P[X = 3] =$ the jump size at $3 = 1 - 0.784 = 0.216 = (6/10)^3$.

In conclusion,

$$\boxed{p_X(x) = \begin{cases} 0.064, & x = 0, \\ 0.288, & x = 1, \\ 0.432, & x = 2, \\ 0.216, & x = 3, \\ 0, & \text{otherwise.} \end{cases}}$$

(b) Among all the pmf that we discussed in class, only one can have support = $\{0, 1, 2, 3\}$ with unequal probabilities. This is the $\boxed{\text{binomial}}$ pmf. To check that it really is Binomial, recall that the pmf for binomial X is given by $p_X(x) = \binom{n}{x}p^x(1-p)^{(n-x)}$ for $x = 0, 1, 2, \ldots, n$. Here, $n = 3$. Furthermore, observe that $p_X(0) = (1-p)^n$. By comparing $p_X(0)$ with what we had in part (a), we have $1 - p = 2/5$ or $p = 3/5$. For $x = 1, 2, 3$, plugging in $p = 3/5$ and $n = 3$ in to $p_X(x) = \binom{n}{x}p^x(1-p)^{(n-x)}$ gives the same values as what we had in part (a). So, X is a binomial RV.

Extra Questions

Here are some questions for those who want extra practice.

Problem 3. When n is large, binomial distribution Binomial(n, p) becomes difficult to compute directly because of the need to calculate factorial terms. In this question, we will consider an approximation when p is close to 0. In such case, the binomial can be approximated[1] by the Poisson distribution with parameter $\alpha = np$.

(a) Let $X \sim$ Binomial$(12, 1/36)$. (For example, roll two dice 12 times and let X be the number of times a double 6 appears.) Evaluate $p_X(x)$ for $x = 0, 1, 2$.

(b) Compare your answers in the previous part with the Poisson approximation.

(c) Compare `MATLAB` plots of $p_X(x)$ and the pmf of $\mathcal{P}(np)$.

(d) In one of the New York state lottery games, a number is chosen at random between 0 and 999. Suppose you play this game 250 times. Use the Poisson approximation to estimate the probability that you will never win and compare this with the exact answer.

Solution:

(a) 0.7132, 0.2445, 0.0384.

(b) 0.7165, 0.2388, 0.0398.

(c) See Figure **??**.

[1] More specifically, suppose X_n has a binomial distribution with parameters n and p_n. If $p_n \to 0$ and $np_n \to \alpha$ as $n \to \infty$, then

$$P[X_n = k] \to e^{-\alpha}\frac{\alpha^k}{k!}.$$

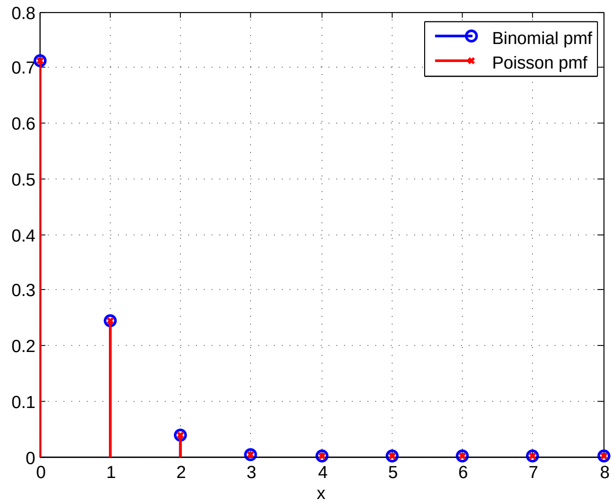

Figure 6.2: Poisson Approximation

(d) [Durrett, 2009, Q2.41] Let W be the number of wins. Then, $W \sim$ Binomial$(250, p)$ where $p = 1/1000$. Hence,

$$P[W = 0] = \binom{250}{0} p^0 (1-p)^{250} \approx 0.7787.$$

If we approximate W by $\Lambda \sim \mathcal{P}(\alpha)$. Then we need to set

$$\alpha = np = \frac{250}{1000} = \frac{1}{4}.$$

In which case,

$$P[\Lambda = 0] = e^{-\alpha}\frac{\alpha^0}{0!} = e^{-\alpha} \approx 0.7788$$

which is very close to the answer from direct calculation.

Problem 4. For each of the following families of random variable X, find the value(s) of x which maximize $p_X(x)$. (This can be interpreted as the "mode" of X.)

(a) $\mathcal{P}(\alpha)$

(b) Binomial(n, p)

(c) $\mathcal{G}_0(\beta)$

(d) $\mathcal{G}_1(\beta)$

Remark [Y&G, p. 66]:

- For statisticians, the mode is the most common number in the collection of observations. There are as many or more numbers with that value than any other value. If there are two or more numbers with this property, the collection of observations is called multimodal. In probability theory, a **mode** of random variable X is a number x_{mode} satisfying
$$p_X\left(x_{\text{mode}}\right) \geq p_X(x) \quad \text{for all } x.$$
- For statisticians, the median is a number in the middle of the set of numbers, in the sense that an equal number of members of the set are below the median and above the median. In probability theory, a median, X_{median}, of random variable X is a number that satisfies
$$P\left[X < X_{\text{median}}\right] = P\left[X > X_{\text{median}}\right].$$
- Neither the mode nor the median of a random variable X need be unique. A random variable can have several modes or medians.

Solution: We first note that when $\alpha > 0$, $p \in (0,1)$, $n \in \mathbb{N}$, and $\beta \in (0,1)$, the above pmf's will be strictly positive for some values of x. Hence, we can discard those x at which $p_X(x) = 0$. The remaining points are all integers. To compare them, we will evaluate $\frac{p_X(i+1)}{p_X(i)}$.

(a) For Poisson pmf, we have
$$\frac{p_X(i+1)}{p_X(i)} = \frac{\frac{e^{-\alpha}\alpha^{i+1}}{(i+1)!}}{\frac{e^{-\alpha}\alpha^{i}}{i!}} = \frac{\alpha}{i+1}.$$

Notice that

- $\frac{p_X(i+1)}{p_X(i)} > 1$ if and only if $i < \alpha - 1$.
- $\frac{p_X(i+1)}{p_X(i)} = 1$ if and only if $i = \alpha - 1$.
- $\frac{p_X(i+1)}{p_X(i)} < 1$ if and only if $i > \alpha - 1$.

Let $\tau = \alpha - 1$. This implies that τ is the place where things change. Moving from i to $i+1$, the probability strictly increases if $i < \tau$. When $i > \tau$, the next probability value (at $i+1$) will decrease.

(i) Suppose $\alpha \in (0,1)$, then $\alpha - 1 < 0$ and hence $i > \alpha - 1$ for all i. (Note that i are are nonnegative integers.) This implies that the pmf is a strictly decreasing function and hence the maximum occurs at the first i which is $i = 0$.

(ii) Suppose $\alpha \in \mathbb{N}$. Then, the pmf will be strictly increasing until we reaches $i = \alpha - 1$. At which point, the next probability value is the same. Then, as we further increase i, the pmf is strictly decreasing. Therefore, the maximum occurs at $\alpha - 1$ and α.

(iii) Suppose $\alpha \notin \mathbb{N}$ and $\alpha \geq 1$. Then we will have have any $i = \alpha - 1$. The pmf will be strictly increasing where the last increase is from $i = \lfloor \alpha - 1 \rfloor$ to $i + 1 = \lfloor \alpha - 1 \rfloor + 1 = \lfloor \alpha \rfloor$. After this, the pmf is strictly decreasing. Hence, the maximum occurs at $\lfloor \alpha \rfloor$.

To summarize,

$$\arg\max_x p_X(x) = \boxed{\begin{cases} 0, & \alpha \in (0,1), \\ \alpha - 1 \text{ and } \alpha, & \alpha \text{ is an integer}, \\ \lfloor \alpha \rfloor, & \alpha > 1 \text{ is not an integer.} \end{cases}}$$

(b) For binomial pmf, we have

$$\frac{p_X(i+1)}{p_X(i)} = \frac{\frac{n!}{(i+1)!(n-i-1)!} p^{i+1} (1-p)^{n-i-1}}{\frac{n!}{i!(n-i)!} p^i (1-p)^{n-i}} = \frac{(n-i)\,p}{(i+1)(1-p)}.$$

Notice that

- $\frac{p_X(i+1)}{p_X(i)} > 1$ if and only if $i < np - 1 + p = (n+1)p - 1$.
- $\frac{p_X(i+1)}{p_X(i)} = 1$ if and only if $i = (n+1)p - 1$.
- $\frac{p_X(i+1)}{p_X(i)} < 1$ if and only if $i > (n+1)p - 1$.

Let $\tau = (n+1)p - 1$. This implies that τ is the place where things change. Moving from i to $i+1$, the probability strictly increases if $i < \tau$. When $i > \tau$, the next probability value (at $i+1$) will decrease.

(i) Suppose $(n+1)p$ is an integer. The pmf will strictly increase as a function of i, and then stays at the same value at $i = \tau = (n+1)p - 1$ and $i + 1 = (n+1)p - 1 + 1 = (n+1)p$. Then, it will strictly decrease. So, the maximum occurs at $(n+1)p - 1$ and $(n+1)p$.

(ii) Suppose $(n+1)p$ is not an integer. Then, there will not be any i that is $= \tau$. Therefore, we only have the pmf strictly increases where the last increase occurs when we goes from $i = \lfloor \tau \rfloor$ to $i + 1 = \lfloor \tau \rfloor + 1$. After this, the probability is strictly decreasing. Hence, the maximum is unique and occur at $\lfloor \tau \rfloor + 1 = \lfloor (n+1)p - 1 \rfloor + 1 = \lfloor (n+1)p \rfloor$.

To summarize,

$$\arg\max_x p_X(x) = \boxed{\begin{cases} (n+1)p - 1 \text{ and } (n+1)p, & (n+1)p \text{ is an integer,} \\ \lfloor (n+1)p \rfloor, & (n+1)p \text{ is not an integer.} \end{cases}}$$

(c) $\frac{p_X(i+1)}{p_X(i)} = \beta < 1$. Hence, $p_X(i)$ is strictly decreasing. The maximum occurs at the smallest value of i which is $\boxed{0}$.

(d) $\frac{p_X(i+1)}{p_X(i)} = \beta < 1$. Hence, $p_X(i)$ is strictly decreasing. The maximum occurs at the smallest value of i which is $\boxed{1}$.