Stat 345/445: Theoretical Statistics I: Homework 3 Solutions

Textbook Exercises

1.45 (445: 2 pts.) Show that the induced probability function defined in (1.4.1) defines a legitimate probability function in that it satisfies the Kolmogorov Axioms.

$\mathcal{X}$ is finite. Therefore $\mathcal{B}$ is the set of all subsets of $\mathcal{X}$.

We must verify each of the three properties in Definition 1.2.4.

1. If $A \in \mathcal{B}$ then $P_X(A) = P(\cup_{x_i \in A}\{s_j \in S : \ X(s_j) = x_i\}) \geq 0$ since P is a probability function.
2. $P_X(\mathcal{X}) = P(\cup_{i=1}^{m}\{s_j \in S : \ X(s_j) = x_i\}) = P(S) = 1$
3. If $A_1, A_2, \cdots \in \mathcal{B}$ and pairwise disjoint then

$$\begin{aligned}
P_X(\cup_{k=1}^{\infty} A_k) &= P(\cup_{k=1}^{\infty}\{\cup_{x_i \in A_k}\{s_j \in S : \ X(s_j) = x_i\}\}) \\
&= \sum_{k=1}^{\infty} P(\cup_{x_i \in A_k}\{s_j \in S : \ X(s_j) = x_i\}) \\
&= \sum_{k=1}^{\infty} P_X(A_k),
\end{aligned}$$

1.49 (445: 2 pts.) A cdf F_X is *stochastically greater* than a cdf F_Y if $F_X(t) \leq F_Y(t)$ for all t and $F_X(t) < F_Y(t)$ for some t. Prove that if $X \sim F_X$ and $Y \sim F_Y$, then

$$P(X > t) \geq P(Y > t) \quad \text{for every } t$$

and

$$P(X > t) > P(Y > t) \quad \text{for some } t$$

that is, X tends to be bigger than Y.

For every t, $\ F_X(t) \leq F_Y(t)$.

Thus we have

$$P(X > t) = 1 - P(X \leq t) = 1 - F_X(t) \geq 1 - F_Y(t) = 1 - P(Y \leq t) = P(Y > t).$$

$$P(X > t) \geq P(Y > t) \quad \text{for every } t$$

And for some t^*, $\ F_X(t^*) < F_Y(t^*)$.

Then we have that

$$P(X > t^*) = 1 - P(X \leq t^*) = 1 - F_X(t^*) > 1 - F_Y(t^*) = 1 - P(Y \leq t^*) = P(Y > t^*).$$

$$P(X > t) > P(Y > t) \quad \text{for some } t$$

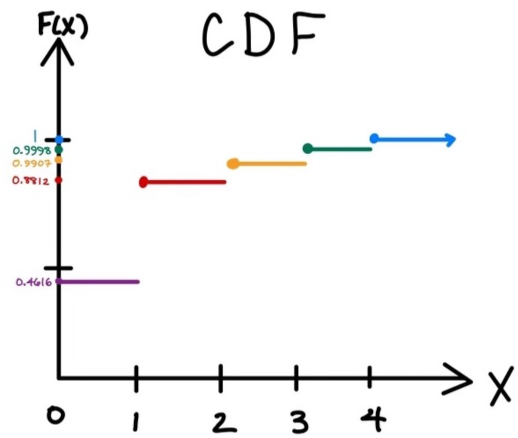
1.51 (345 & 445: 2 pts.) An appliance store receives a shipment of 30 microwave ovens, 5 of which are (unknown to the manager) defective. The store manager selects 4 oven at random, without replacement, and tests to see if they are defective. Let X = number of defectives found. Calculate the pmf and cdf of X and plot the cdf.

This kind of random variable is called hypergeometric in Chapter 3. The probabilities are obtained by counting arguments, as follows

$$P(X=x) = \frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}}$$

The cdf is a step function with jumps at $x = 0, 1, 2, 3,$ and 4.

x	$f_X(x) = P(X=x)$	$F_X(x)$
0	$\binom{5}{0}\binom{25}{4} \Big/ \binom{30}{4} \approx 0.4616$	0.4616
1	$\binom{5}{1}\binom{25}{3} \Big/ \binom{30}{4} \approx 0.4196$	0.8812
2	$\binom{5}{2}\binom{25}{2} \Big/ \binom{30}{4} \approx 0.1095$	0.9907
3	$\binom{5}{3}\binom{25}{1} \Big/ \binom{30}{4} \approx 0.0091$	0.9998
4	$\binom{5}{4}\binom{25}{0} \Big/ \binom{30}{4} \approx 0.0002$	1



1.52 (345 & 445: 2 pts.) Let X be a continuous random variable with pdf $f(x)$ and cdf $F(x)$. For a fixed number x_0, define the function

$$g(x) = \begin{cases} f(x)/[1-F(x_0)] & x \geq x_0 \\ 0 & x < x_0. \end{cases}$$

Prove that $g(x)$ is a pdf. (Assume that $F(x_0) < 1$.)

The function $g(\cdot)$ is clearly positive. Also,

$$\int_{x_0}^{\infty} g(x)dx = \int_{x_0}^{\infty} \frac{f(x)}{1-F(x_0)}dx = \frac{F(x)|_{x_0}^{\infty}}{1-F(x_0)} = \frac{1-F(x_0)}{1-F(x_0)} = 1.$$

Thus, $g(x)$ is a pdf.

1.53 (345 & 445: 2 pts.) A certain river floods every year. Suppose that the low-water mark is set at 1 and the high-water mark Y has distribution function

$$F_Y(y) = P(Y \leq y) = 1 - \frac{1}{y^2}, \quad 1 \leq y < \infty.$$

(a) Verify that $F_Y(y)$ is a cdf.

- $\lim_{y\to-\infty} F_Y(y) = \lim_{y\to-\infty} 0 = 0$
- $\lim_{y\to\infty} F_Y(y) = \lim_{y\to\infty} 1 - \frac{1}{y^2} = 1$
- $\begin{cases} \text{For } y \leq 1, & F_Y(y) = 0 \text{ is constant} \\ \text{For } y > 1, & \frac{d}{dy}F_Y(y) = \frac{2}{y^3} > 0, \text{ so } F_Y \text{ is increasing.} \end{cases}$ For all y, F_Y is nondecreasing.

Therefore F_Y is a cdf.

(b) Find $f_Y(y)$, the pdf of Y.

The pdf is $f_Y(y) = \frac{d}{dy}F_Y(y) = \begin{cases} \frac{2}{y^3} & \text{if } y > 1 \\ 0 & \text{if } y \leq 1. \end{cases}$

(c) If the low-water mark is reset at 0 and we use a unit of measurement that is $\frac{1}{10}$ of that given previously, the high-water mark becomes $Z = 10(Y-1)$. Find $F_Z(z)$.

$$F_Z(z) = P(Z \leq z) = P(10(Y-1) \leq z) = P(Y \leq (z/10)+1) = F_Y((z/10)+1).$$

Thus,

$$F_Z(z) = \begin{cases} 0 & \text{if } z \leq 0 \\ 1 - \left(\frac{1}{(\frac{z}{10}+1)^2}\right) & \text{if } z > 0. \end{cases}$$

Extra Problems

1. (345: 2 pts.) Let $f(x)$ and $g(x)$ be two probability density functions (pdfs). Show that $\alpha f(x) + (1-\alpha)g(x)$ is a pdf, where $0 \leq \alpha \leq 1$.

$f(x)$ and $g(x)$ are two pdfs.

$$f(x) \geq 0, \quad g(x) \geq 0, \quad \int_{-\infty}^{\infty} f(x)dx = 1, \quad \int_{-\infty}^{\infty} g(x)dx = 1$$

Let $h(x) = \alpha f(x) + (1-\alpha)g(x), \quad 0 \leq \alpha \leq 1$

We have $h(x) \geq 0$ as $f(x) \geq 0, g(x) \geq 0$ and $0 \leq \alpha \leq 1$

$$\int_{-\infty}^{\infty} h(x)dx = \int_{-\infty}^{\infty} \alpha f(x) + (1-\alpha)g(x)dx = \alpha \int_{-\infty}^{\infty} f(x)dx + (1-\alpha)\int_{-\infty}^{\infty} g(x)dx = \alpha + (1-\alpha) = 1$$

Thus, $\alpha f(x) + (1-\alpha)g(x)$ is a pdf.

2. (345: 2 pts.) Suppose that X has the pdf $f(x) = cx^2$ for $0 \leq x \leq 1$ and $f(x) = 0$ otherwise.

(a) Find c.

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^1 cx^2 dx = \frac{1}{3}cx^3|_0^1 = \frac{1}{3}c = 1 \implies c = 3$$

(b) Find the cdf.

$$f(x) = \begin{cases} 3x^2 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\int_{-\infty}^{x} f(x)dx = \int_0^x 3x^2 dx = x^3|_0^x = x^3 \quad 0 \leq x \leq 1$$

$$F(x) = \begin{cases} 0 & x < 0 \\ x^3 & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

(c) What is $P(0.1 \leq X < 0.5)$?

$$P(0.1 \leq x < 0.5) = F(0.5) - F(0.1) = (0.5)^3 - (0.1)^3 = 0.124$$