Probability 1 - Homework 5

Philip Warton

December 3, 2020

Problem 1

The density function of X is given by

$$f(x) = \begin{cases} c/x^4, & x \geqslant 1 \\ 0, & \text{otherwise} \end{cases}$$

(i)

Show that $c = 3$.

Proof. Since f is a probability denisty function, it must be the case that $\int_{\mathbb{R}} f = 1$. Using this fact, we can show that c must be equal to 3. We begin by integrating the function on the real line.

$$\begin{aligned} \int_{\mathbb{R}} f(x)dx &= \int_{-\infty}^{1} f(x)dx + \int_{1}^{\infty} f(x)dx \\ &= \int_{-\infty}^{1} 0dx + \int_{1}^{\infty} cx^{-4}dx \\ &= \left[-\frac{1}{3}cx^{-3}\right]_1^\infty \\ &= \lim \left[-\frac{1}{3}cn^{-3}\right] + \frac{1}{3}c1^{-3} \\ &= 0 + \frac{1}{3}c \end{aligned}$$

Since we know that the total probability must be equal to 1, it follows that $c = 3$. $\square$

(ii)

Compute $E[X]$ and $Var(X)$.

Firstly we can compute $E[X]$. We know that this is equal to $\int_{\mathbb{R}} xf(x)dx = \int_1^\infty x3x^{-4}dx = \int_1^\infty 3x^{-3}dx$. This integral evaluates to $\left[-\frac{3}{2}x^{-2}\right]_1^\infty = \lim\left[-\frac{3}{2}n^{-2}\right] - \left[-\frac{3}{2}1^{-2}\right] = \frac{3}{2}$. Then to compute the variance, we first wish to compute $E[X^2]$. We simply write

$$E[X^2] = \int_{\mathbb{R}} x^2 f(x)dx = \int_1^\infty x^2 3x^{-4}dx = \int_1^\infty 3x^{-2}dx = \left[-3x^{-1}\right]_1^\infty = 0 + 3$$

Then having computed both $E[X]$ and $E[X^2]$ we know that we can compute the variance using $Var(X) = E[X^2] - E[X]^2$. This gives us $Var(X) = 3 - \frac{3^2}{2^2} = \frac{3}{4}$.

Problem 2

Let Z be a standard normal random variable $Z \sim N(0,1)$. Compute $E[e^Z]$.

We know that for some function we can simply integrate our density multiplied by our function to compute the expectation. So we say that $E[e^Z] = \int_{\mathbb{R}} e^x \sqrt{2\pi}^{-1} e^{-\frac{x^2}{2}}dx = \int_{-\infty}^{\infty} \sqrt{2\pi}^{-1} e^{x-\frac{x^2}{2}}dx$. We can first complete the square on the term $x - \frac{x^2}{2}$ which gives us $x - \frac{x^2}{2} = (1/2)((x-1)^2+1)$. Then let $u = x - 1$. By exchanging variables, we still have positive and negative infinite bounds, and we write $E[e^Z] = \int_{\mathbb{R}} \sqrt{2\pi}^{-1} e^{(1/2)(u^2+1)}du$. Now we can factor out our $\sqrt{2\pi}^{-1}$ and we can also factor out $e^{\frac{1}{2}}$. This gives us

$\frac{\sqrt{e}}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{u^2/2}du$. Then we exchange variables again, defining $v = u/\sqrt{2}$, so that $dv = du/\sqrt{2}$. Our bounds remain unchanged as they are only scaled by some constant, and we can say

$$\frac{\sqrt{2e}}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{v^2}dv = \frac{\sqrt{2e}}{\sqrt{2\pi}}\sqrt{2\pi} = \sqrt{2e}$$

Finally $E[e^Z] = \sqrt{2e}$.

Problem 3

Show that $Var(X+Y) = Var(X) + Var(Y)$ for two independent random variables X and Y.

Proof. We should first note that $E[X]E[Y] = E[XY]$, and also that $Var(Z) = E[Z^2] + E[Z]^2$. Then we can simply use algebraic manipulations, to show that this equality holds.

$$\begin{aligned}
Var(X+Y) &= E[(X+Y)^2] - E[X+Y]^2 \\
&= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 \\
&= E[X^2] + 2E[XY] + E[Y^2] - (E[X]^2 + 2E[X]E[Y] + E[Y^2]) \\
&= E[X^2] - E[X]^2 + E[Y^2] - E[Y]^2 + 2E[XY] - 2E[XY] \\
&= Var(X) + Var(Y)
\end{aligned}$$

And thus we have shown that variance is distributive over addition of independent random variables. $\square$

Problem 4

Let X_1, X_2, X_3 be uniform random variables on $[0,1]$. Then find the density of $Y = X_1 + X_2 + X_3$.

Of course we will compute the cumulative distribution function $F(Y)$, and then differentiate it in order to get the density. We can imagine values of Y as points that lie within the unit cube in $\mathbb{R}^3$. If $Y = a$ for some $a \in [0,3]$, then we know that y is a point in the intersection of the plane $a = x_1 + x_2 + x_3$ and the unit cube. So for the cumulative distribution function, we say that $P(Y < a)$ is the volume of everything underneath this plane $a = x_1 + x_2 + x_3$ (a plane with a unit vector as a normal, offset from the origin by a) intersected with the unit cube. For $a \in [0,1]$, the shape will be a right corner piece of the unit cube which we can compute the volume of as

$$\int_0^a \int_0^{x_3} \int_0^{x_2} 1 dx_1 dx_2 dx_3 = \int_0^a \int_0^{x_3} x_2 dx_2 dx_3 = \int_0^a \frac{x_3^2}{2} dx_3 = \frac{a^3}{6}$$

Then for $a \in [1,2]$, we have an odd shape. We could describe it as this right corner piece truncated at its spikes that go beyond 1. The pieces that become truncated are 3 corners that are identical to these right triangle corner pieces except with length $a-1$. So the volume of each of these will be $\frac{(a-1)^3}{6}$. So then for $a \in [1,2]$ we say $P(Y < a) = \frac{a^3}{6} - 3\frac{a-1)^3}{6}$. Then finally for $a \in [2,3]$ we have the unit cube minus the volume of a right triangle corner piece with side length $3-a$, i.e. we have the volume $1 - \frac{(3-a)^3}{6}$. Then for $a < 0, P(Y < a) = 0$ and for $a > 3, P(Y < a) = 1$. So we have the cumulative distribution function

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ \frac{y^3}{6}, & 0 \leqslant y < 1 \\ \frac{y^3}{6} - 3\frac{(y-1)^3}{6}, & 1 \leqslant y < 2 \\ 1 - \frac{(3-y)^3}{6}, & 2 \leqslant y < 3 \\ 1, & y \geqslant 3 \end{cases}$$

Then, by differentiating this function, we get our density function $f_y(y)$.

$$f_y(y) = \begin{cases} \frac{y^2}{2}, & 0 \leqslant y < 1 \\ \frac{y^2}{2} - (y-1)^2, & 1 \leqslant y < 2 \\ \frac{(3-y)^2}{2}, & 2 \leqslant y < 3 \\ 0, & \text{otherwise} \end{cases}$$

Problem 5

The sample size $n = 1210$, with $p = \frac{1}{11}$. Then we will compute $P\{97.5 \leqslant \text{ the number of successes } \leqslant 116.5\}$. Then we write $\frac{97.5-(1210/11)}{\sqrt{12100/121}} = -\frac{5}{4}$. Then to normalize our upper bound we write $\frac{116.5-(1210/11)}{\sqrt{12100/121}} = \frac{13}{20} = .65$.

$P\{97.5 \leqslant \text{ the number of successes } \leqslant 116.5\} = P(-1.25 < Z < .65) = 1-P(Z > .65)-P(Z > 1.25) \approx 1-.2578-.1056 = .6366$

Problem 6

The sample size is $n = 90000$, $p = \frac{1}{2}$. Then we want

$$P\{45032 \leqslant \text{ the number of heads } \leqslant 45069\}$$

Our lower bound will be $\frac{45032-45000}{\sqrt{22500}} = \frac{16}{75} \approx .2133$. Then our upper bound will be $\frac{45069-45000}{\sqrt{22500}} = \frac{23}{50} \approx .46$. Finally we can compute

$P\{45032 \leqslant \text{ the number of heads } \leqslant 45069\} \approx P(.2133 < Z < .46) = 1-P(Z > .2133)-P(Z > .46) \approx 1-.4168-.3228 = .2604$

Problem 7

With a sample size $n = 18000$ and a probability $p = \frac{1}{6}$, compute the probability of getting at least 3060 successful trials, $P\{3059.5 < \text{ number of successes }\}$. Then we get our lower bound and consult the Z-table. We say that $\frac{3059.5-3000}{\sqrt{2500}} = 1.19$. Then

$$P(Z > 1.19) \approx .1170$$