Math 340 HW 7

Asa Royal (ajr74) [collaborators: none]

March 22 2024

1. Since for a random walk $S_n = X_1 + \ldots + X_n$ where X_1 are iid random variables, we can apply the central limit theorem to find the probability that S_n lies within some range. In this problem, we are exploring

$$\lim_{n\to\infty} \mathbb{P}(|S_n| < n^r) = 0 \tag{1}$$

That expression is equivalent to

$$\mathbb{P}(-n^r \leq X_1 + \ldots + X_n \leq n^r) \tag{2}$$

We can make this look like the CLT by subracting μn from all terms in the inequality, then dividing by $\sqrt{n\sigma^2}$. Thus, (1) and (2) are equivalent to

$$\mathbb{P}\left(\frac{-n^r - \mu n}{\sqrt{n\sigma^2}} \leq \frac{X_1 + \ldots + X_n - \mu n}{\sqrt{n\sigma^2}} \leq \frac{n^r - \mu n}{\sqrt{n\sigma^2}}\right) \tag{3}$$

We can easily calculate $\sigma = \text{Var}(X_i)$. Since X_i takes values 0 and 1 with equal probability, $\mathbb{E}[X_i] = 0$ and:

$$\begin{aligned} \text{Var}(X_i) &= \mathbb{E}[{X_i}^2] - \mathbb{E}[X_i]^2 && (1)\\ &= (-1)^2(1/2) + (1)^2(1/2) && (2)\\ &= 1 && (3) \end{aligned}$$

Returning to (3) and noting $\mu n = 0$, we can now express

$$\lim_{n\to\infty} \mathbb{P}(|S_n| < n^r) = \mathbb{P}\left(\frac{-n^r}{\sqrt{n}} \leq \frac{X_1 + \ldots + X_n - \mu n}{\sqrt{n\sigma^2}} \leq \frac{n^r}{\sqrt{n}}\right)$$

Per the CLT, this is

$$\int_{-n^{\left(r-\frac{1}{2}\right)}}^{n^{\left(r-\frac{1}{2}\right)}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy$$

If $r > 1/2$, the upper bound of the integral tends to infinity while the lower bound tends to negative infinity, so the expression becomes $\int_{-\infty}^{\infty} f(x)dx = 1$. On the other hand, if $r < 1/2$, the upper bound will evaluate to n^{-x} for $x \in (-1/2, 0)$. As $n \to \infty$, this value tends toward zero. The lower bound tends towards $-n^{-x}$, which is also 0. Thus, the CLT integral becomes $\int_0^0 f(x)dx = 0$.

$\boxed{\text{In conclusion, if } \alpha = 1/2\text{, both of the statements expressed in the problem are true.}}$

2. Let $X_1, \ldots, X_n$ be random variables representing the weight of the individual boxes. We wish to estimate $\mathbb{P}(4850 \leq X_1 + \ldots + X_n \leq 5150$. Since $X_1, \ldots, X_n$ are iid random variables and n is relatively large, the sum of the variables roughly follows a normal distribution. We thus make use of the central limit theorem:

$$\begin{aligned} \mathbb{P}(4850 \leq X_1 + \ldots + X_n \leq 5150) &= \mathbb{P}\left(\frac{4850 - \mu n}{\sqrt{n\sigma^2}} \leq \frac{X_1 + \ldots + X_n - \mu n}{\sqrt{n\sigma^2}} \leq \frac{5150 - \mu n}{\sqrt{n\sigma^2}}\right) && (1)\\ &= \mathbb{P}\left(\frac{4850 - 5000}{\sqrt{2500(2)^2}} \leq \frac{X_1 + \ldots + X_n - \mu n}{\sqrt{n\sigma^2}} \leq \frac{5150 - 5000}{\sqrt{2500(2)^2}}\right) && (2)\\ &= \mathbb{P}\left(-150 \leq \frac{X_1 + \ldots + X_n - \mu n}{\sqrt{n\sigma^2}} \leq 150\right) && (3)\\ & && (4) \end{aligned}$$

Per the CLT, (3) is equivalent to

$$\int_{-150}^{150} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy \tag{5}$$

Which, in turn, is equal to $\boxed{\Phi(150) - \Phi(-150)}$.

3. Consider the tail of Y.

$$\begin{aligned}\mathbb{P}(Y > t) &= \mathbb{P}(cX > t) \\ &= \mathbb{P}\left(X > \frac{t}{c}\right)\end{aligned}$$

Since $X \sim \text{Exponential}(\lambda)$, this is

$$\mathbb{P}\left(X > \frac{t}{c}\right) = e^{-\frac{\lambda}{c}t}$$

Y's tail probability is equivalent to that of an exponentially-distributed r.v. with parameter $-\frac{\lambda}{c}$. Thus $Y \sim$ Exponential $\left(-\frac{\lambda}{c}\right)$.

4. **True**. If $f(x) < 0$ for some $x = a$, $F_x = \int f(x)dx$ would decrease along some interval $[a-\varepsilon, a+\varepsilon]$, violating the definition of a CDF. If $f(x) > 1$ for some $x = \beta$, then $\int_{\beta-\varepsilon}^{\beta+\varepsilon} f(x)dx > 1$. Thus, clearly $\int_{-\infty}^{\infty} f(x)dx \neq 1$, violaiting the definition of a CDF.

5. Suppose that X is uniformly distributed on the interval $[2, 4]$.

   (i) What is the density for X?

   $$f(x) = \begin{cases} 1/2 & \text{if } x \in [2,4] \\ 0 & \text{otherwise} \end{cases}$$

   (ii) What is the CDF for X?

   $$F_x = \begin{cases} 0 & \text{if } x \leq 2 \\ \frac{1}{2}(x-2) & \text{if } x \in [2,4] \\ 1 & \text{if } x \geq 4 \end{cases}$$

   (iii) What are the CDF and density for the random variable $Y = X^2 + 1$?

   i. Find F_Y:

   $$\begin{aligned}\mathbb{P}(Y \leq b) &= \mathbb{P}(X^2 + 1 \leq b) \\ &= \mathbb{P}(X^2 \leq b - 1) \\ &= \mathbb{P}(X \leq \sqrt{b-1} \\ &= \mathbb{P}(X \in [2, \sqrt{b-1}]) \\ &= \frac{\sqrt{b-1} - 2}{2}\end{aligned}$$

   But note that the Y has 0 density when $X \leq 2, Y \leq 5$. Thus,

   $$F_Y = \begin{cases} 0 & \text{if } Y \leq 5 \\ \frac{\sqrt{b-1}-2}{2} & \text{if } 5 \leq b \leq 17 \\ 1 & \text{if } Y \geq 17 \end{cases}$$

   ii. Find density of Y:

   $$\begin{aligned}f(b) &= \frac{d}{db}\left(\frac{\sqrt{b-1}-2}{2}\right) \\ &= \frac{1}{2}\frac{d}{db}(\sqrt{b-1} - 2) \\ &= \frac{1}{2}\left(\frac{1}{2\sqrt{b-1}}\right) \\ &= \frac{1}{4\sqrt{b-1}}\end{aligned}$$

Taking into account the zero density regions,

$$\begin{cases} 0 & \text{if } Y \leq 5 \\ \frac{1}{4\sqrt{b-1}} & \text{if } 5 \leq Y \leq 17 \\ 0 & \text{if } Y \geq 17 \end{cases}$$

(iv) What are the mean and variance of Y?

i. Find the mean of Y:

Per the linearity of expectation,

$$\mathbb{E}[Y] = \mathbb{E}[X^2 + 1] = \mathbb{E}[X^2] + 1 \tag{1}$$

We find $\mathbb{E}[X^2]$ below:

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 f(x) dx && && (1) \\
&= \int_2^4 x^2 f(x) dx && \text{limit domain to interval} && (2) \\
&= \frac{1}{2} \int_2^4 x^2 && \text{plug in uniform dens.} && (3) \\
&= \frac{1}{6} x^3 \Big|_2^4 && && (4) \\
&= 28/3 && && (5)
\end{aligned}$$

Combining (1) and (5), we conclude $\boxed{\mathbb{E}[Y] = 31/3}$.

ii. Find the variance of Y:

$\mathrm{Var}(Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2$. We calculated $\mathbb{E}[Y]$ above. We calculate $\mathbb{E}[Y^2]$ below, applying the linearity of expectation and the formula for the expectation of a function of a continuous r.v.

$$\begin{aligned}
\mathrm{Var}(Y) &= \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \\
&= \mathbb{E}[X^4 + 2x^2 + 1] - (31/3)^2 \\
&= \mathbb{E}[X^4] + \mathbb{E}[2X^2] + 1 - (31/3)^2 \\
&= \int_2^4 \frac{1}{2} x^4 dx + 2\mathbb{E}[X^2] + 1 - (31/3)^2 && \mathbb{E} \text{ of cont rv} \\
&= \frac{1}{10} x^5 \Big|_2^4 + 2(28/3) + 1 - (31/3)^2 && \text{sub } \mathbb{E}[X^2] \text{ from (5)} \\
&= \frac{1}{10}(1024 - 32) + 56/3 + 1 - (32/3)^2 \\
&= 992/10 + 56/3 + 1 - (32/3)^2 \\
&= \boxed{229/45}
\end{aligned}$$

6. ..

(a) Find c:

For any pdf $f(x)$, $\int_{-\infty}^{\infty} f(x) = 1$, which means in this case, $\int_{-1}^{1} cx^2 = 1$. Solving for c:

$$\begin{aligned}
\int_{-1}^{1} cx^2 dx &= 1 \\
\frac{c}{3} x^3 \Big|_{-1}^{1} &= 1 \\
\frac{2c}{3} &= 1 \\
&\boxed{c = 3/2}
\end{aligned}$$

(b) Find $\mathbb{P}(X > 1/2)$

$$
\begin{aligned}
\mathbb{P}(X > 1/2) &= 1 - \mathbb{P}(X \leq 1/2) && \text{express tail in terms of CDF} \\
&= 1 - \int_{-\infty}^{1/2} f(x) \\
&= 1 - \left(\int_{-\infty}^{-1} f(x)dx + \int_{-1}^{1/2} f(x)dx \right) \\
&= 1 - \int_{-1}^{1/2} \frac{3}{2}x^2 dx && \text{since density is 0 outside } [-1, 1] \\
&= 1 - \frac{1}{2}x^3\Big|_{-1}^{1/2} \\
&= 1 - \left(\frac{1}{2}\right)\left(\frac{9}{8}\right) \\
&\boxed{= 7/16}
\end{aligned}
$$

7. We will show that $F_Y(t) = t$, revealing that $Y \sim$ Unif.

$$
\begin{aligned}
F_Y(t) &= \mathbb{P}(Y \leq t) && \text{def of CDF} && (1) \\
&= \mathbb{P}(F(X) \leq t) && && (2) \\
&= \mathbb{P}\left(X \leq F^{-1}(t)\right) \text{ on } 0 \leq t \leq 1 && \text{ineq. holds b/c F is increasing} && (3) \\
&= F(F^{-1}(t)) \text{ on } 0 \leq t \leq 1 && \text{def of CDF} && (4) \\
&= t \text{ on } 0 \leq t \leq 1 && && (5)
\end{aligned}
$$

Since $F_Y(t) = t$ on $0 \leq t \leq 1$, $\boxed{Y \sim \text{Unif}(0, 1)}$.