Bradley Scott
Data601
HW12

Problem 1

(a)

let X_n be the sum of each days gains or losses

We know $-400k \leq X_n \leq 400k$

$$\mathrm{X_n} = X_1 + X_2 + \cdots + X_{400}$$

$$E[X_n] = nE[X_i] = n * 0 = 0$$

$$var[X_i] = E[X_i^2] = \left(\frac{1}{2}\right)(k^2) + \left(\frac{1}{2}\right)(-k)^2 = k^2$$

$$var[X_n] = nvar[X_i] = 400k^2$$

$$P[-10{,}000 \leq X_n \leq 10{,}000] = \frac{1}{3}$$

We have the formula

$$P\left[a \leq \frac{S_n - E[S_n]}{\sqrt{var[S_n]}} \leq b\right] \to \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}\,dx \; as \; n \to \infty$$

Assume n is large enough to use this approximation

We have $P\left[-a \leq \frac{X_n - 0}{20k} \leq a\right] \to \frac{1}{\sqrt{2\pi}}\int_{-a}^{a} e^{-\frac{x^2}{2}}dx = \frac{1}{3}$

Lets find a that satisfies the above equation (I also verified with this website https://onlinestatbook.com/2/calculators/normal_dist.html)

wolfram alpha estimate that $a \approx 0.43$

$$P\left[-0.43 \leq \frac{X_n}{20k} \leq 0.43\right] \to \frac{1}{3}$$

$$\Rightarrow \; P[-8.6k \leq X_n \leq 8.6k] \to \frac{1}{3}$$

$$\Rightarrow 8.6k = 10{,}000$$
$$\Rightarrow k \approx 1162$$

(b)

want $P[X_n \geq 20{,}000] = \; ?$

with $n = 400$

$$P[X_n \geq 20{,}000] = P[20{,}000 \leq X_n \leq \infty] = P\left[\frac{20{,}000}{20 * 8.6} \leq \frac{X_n}{20k} \leq \infty\right]$$
$$\rightarrow \frac{1}{\sqrt{2\pi}} \int_{116}^{\infty} e^{-\frac{x^2}{2}}\, dx \; as\; n \rightarrow \infty$$

assuming n is large enough that we can use this approximation

and we get 0

To double check out answers we can use the following:

$$P[\,|\frac{S_n - E[S_n]}{n}| \geq t] = P\left[\left(\frac{S_n - E[S_n]}{n}\right)^2 \geq t^2\right] \leq \frac{var[S_n]}{n^2 t^2} = \frac{n\, var[S_1]}{n^2 t^2} = \frac{var[S_1]}{n t^2}$$

$$and\; \frac{var[S_1]}{n t^2} \rightarrow 0 \; as\; n \rightarrow \infty$$

So then

$$P[X_n \geq 20{,}000] = P\left[\frac{X_n - 0}{400} \geq 50\right] \leq \frac{(1162)^2}{400(50)^2} \leq 1.3 \; so\; that's\; not\; helpful$$

Problem 2

$$X_i \sim \exp(\lambda = 1)$$

$$E[X_i] = \lambda = 1$$

$$var[X_i] = \frac{1}{\lambda^2} = 1$$

$$X_{500} = X_1 + X_2 + \cdots + X_{500}$$

$$E[X_{500}] = nE[X_i] = 500$$

$$var[X_{500}] = nvar[X_i] = 500$$

want $P[X_{500} \leq 7 * 6 = 420]$

Normalizing we get $P\left[\frac{X_{500} - 500}{\sqrt{500}} \leq \frac{420 - 500}{\sqrt{500}}\right] = P[Z \leq -3.58]$ *where Z is a normal distribution*

The probability that $Z \leq -3.58$ is off the charts small and is ≈ 0.00017.

Problem 3

$X_1,\ X_2,\ X_3\ and\ X_4 \sim$ random sample from $N(\alpha, \sigma^2)$

$$Y = X_1 \ - 2X_2 + 3X_3 \ - 4X_4$$

We know that a linear combination of independent normal random variables will also be normal so we just need to calculate the mean and variance of Y.

$$E[Y] = E[X_1 \ - 2X_2 + 3X_3 \ - 4X_4] = \alpha - 2\alpha + 3\alpha - 4\alpha = -2\alpha$$

$$var[Y] = var[X_1 \ - 2X_2 + 3X_3 \ - 4X_4] = \sigma^2 + 4\sigma^2 + 9\sigma^2 + 16\sigma^2 = 30\sigma^2$$

So $Y \sim N(-2\alpha,\ 30\sigma^2)$

Problem 4

$X_1,\ X_2,\ X_3,\ X_4,\ X_5 \sim$ random sample from uniform distribution on $[0, \theta]$

estimator $x_{max} = k\ * \max\ (X_1,\ \ldots,\ X_5)$

(a) Find the density of x_{max}

$$X_i \sim f_{X_i}(x) = \{ \begin{matrix} \frac{1}{\theta},\ for\ x \in [0, \theta] \\ 0,\ otherwise \end{matrix}$$

$$X_i \sim F_{X_i}(x) = \{ \begin{matrix} \frac{x}{\theta},\ for\ x \geq 0 \\ 0,\ otherwise \end{matrix}$$

$F_{x_{max}}(X_1,\ \ldots,\ X_5) = P[k\ * \max(X_1,\ \ldots,\ X_5) \leq \mathrm{x}] = P[\max(X_1,\ \ldots,\ X_5) \leq \frac{\mathrm{x}}{\mathrm{k}}]$

This means that for some $i \in \{1,2,3,4,5\}$ there is an $X_i = \frac{x}{k}$ since we're looking at the max

$$= P\left[\left(X_1 \leq \frac{x}{K}\right) \cap \left(X_2 \leq \frac{x}{K}\right) \cap \left(X_3 \leq \frac{x}{K}\right) \cap \left(X_4 \leq \frac{x}{K}\right) \cap \left(X_5 \leq \frac{x}{K}\right)\right]$$

We know that $X_1,\ \ldots,\ X_5$ are independent so we have

$$= P\left[\left(X_i \leq \frac{x}{K}\right)^5\right] = \left(P\left[X_i \leq \frac{x}{k}\right]\right)^5$$

Now if we look at the formula for $F_{X_i}(x)$ we see that $P\left[X_i \leq \frac{x}{k}\right] = F_{X_i}\left(\frac{x}{k}\right)$

$$= \left(F_{X_i}\left(\frac{x}{k}\right)\right)^5 = \left(\frac{\frac{x}{k}}{\theta}\right)^5 = \frac{x^5}{k^5\theta^5}$$

so we have $F_{x_{max}}(X_1, \ \dots, \ X_5) = \frac{x^5}{k^5\theta^5}$

Now if we want $f_{x_{max}}(X_1, \ \dots, \ X_5)$ we can take the partial derivative of $F_{x_{max}}(X_1, \ \dots, \ X_5)$ with respect to X

$$f_{x_{max}}(X_1, \ \dots, \ X_5) = \frac{\partial}{\partial x}\frac{x^5}{k^5\theta^5} = \frac{5x^4}{k^5\theta^5}$$

$$f_{x_{max}}(X_1, \ \dots, \ X_5) = \begin{cases} \frac{5x^4}{k^5\theta^5}, \ x \in [0, k\theta] \\ 0, \ otherwise \end{cases}$$

(b)

$$E[X_{max}] = E[k \ * \max(X_1, \dots, X_5)] = k \ * E[\max(X_1, \ \dots, \ X_5)]$$
$$= k \ * \ \int_0^{k\theta} \max(X_1, \ \dots, \ X_5) \ * \ \frac{5x^4}{k^5\theta^5} dx$$

Note that $\max(X_1, \ \dots, \ X_5) = \mathrm{x}$ for some x so we can replace that in our integral

$= \frac{5}{k^4\theta^5}\int_0^{k\theta} x^5 dx = \frac{5}{k^4\theta^5} \ * [\ \frac{x^6}{6}\Big|_0^{k\theta} = \frac{5k^6\theta^6}{6k^4\theta^5} = \frac{5k^2\theta}{6}$

In order for X_{max} to be an unbiased estimator $E[X_{max}] = \theta$

$$\frac{5k^2\theta}{6} = \theta$$

$$\Rightarrow k^2 = \frac{6}{5}$$

$$\Rightarrow k = \ \pm \sqrt{\frac{6}{5}}$$