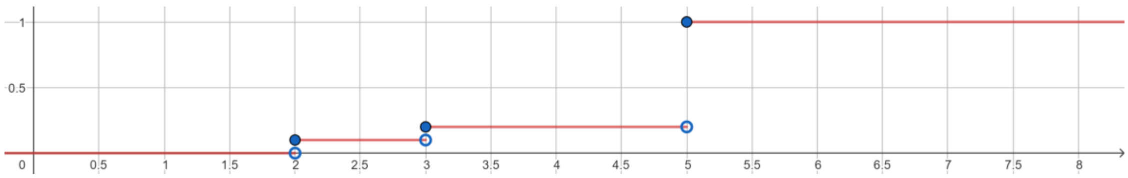
ECO394D Probability and Statistics Homework 3

Sibo Ding

Summer 2023

Problem 1

Part A



$$P(2 < X \leq 4.5) = P(X = 3) = \frac{1}{10}$$

$$P(2 \leq X < 4.5) = P(X = 2) + P(X = 3) = \frac{2}{10}$$

Part B

i. As $X \geq 0$, $P(X^2 \leq 0.25) = P(X \leq 0.5) = \frac{1}{2}$

ii. If $a < 0$, $P(X^2 \leq a) = 0$
If $0 \leq a \leq 1$, $P(X^2 \leq a) = P\left(X \leq \sqrt{a}\right) = \sqrt{a}$
If $a > 1$, $P(X^2 \leq a) = 1$

iii. $f(y) = \frac{d}{dy}\sqrt{y} = \frac{1}{2\sqrt{y}}$ for $0 \leq y \leq 1$
0 otherwise

iv. $E(Y) = \int_0^1 y\, f(y)\, dy$

$$= \int_0^1 \frac{y}{2\sqrt{y}}\, dy$$

$$= \int_0^1 \frac{\sqrt{y}}{2}\, dy$$

$$= \left[\frac{1}{3}x^{\frac{3}{2}}\right]_0^1$$

$$= \frac{1}{3}$$

$$Var(Y) = \int_0^1 (y - \mu)^2\, f(y)\, dy$$

$$
\begin{aligned}
&= \int_0^1 \frac{\left(y-\frac{1}{3}\right)^2}{2\sqrt{y}}\, dy \\
&= \int_0^1 \frac{1}{2}y^{\frac{3}{2}} - \frac{1}{3}y^{\frac{1}{2}} + \frac{1}{18}y^{-\frac{1}{2}}\, dy \\
&= \left[\frac{1}{5}y^{\frac{5}{2}} - \frac{2}{9}y^{\frac{3}{2}} + \frac{1}{9}y^{\frac{1}{2}}\right]_0^1 \\
&= \frac{4}{45}
\end{aligned}
$$

Problem 2

Part A

$$
\begin{aligned}
&\because Z \sim i.i.d.\ N(0,1) \\
&\therefore E(X) = E(Z_1^2 + \cdots + Z_d^2) = d \cdot E(Z^2) \\
&\because E(Z^2) = Var(Z) + [E(Z)^2] = 1 + 0 = 1 \\
&\therefore E(X) = d \cdot E(Z^2) = d
\end{aligned}
$$

$$
\begin{aligned}
&\because Z \sim i.i.d.\ N(0,1) \\
&\therefore Var(X) = Var(Z_1^2 + \cdots + Z_d^2) = d \cdot Var(Z^2) \\
&\because Var(Z^2) = E(Z^4) - [E(Z^2)^2] = 3\sigma^4 - \sigma^4 = 2 \\
&\therefore Var(X) = d \cdot Var(Z^2) = 2d
\end{aligned}
$$

https://math.stackexchange.com/questions/620045/mean-and-variance-of-squared-gaussian-y-x2-where-x-sim-mathcaln0-sigma

Part B

Markov's calculation on expected speed is correct, but his calculation on expected time is wrong. The expected time is not "distance divided by expected speed". It is calculated from the time of each event, like below:

$$
\begin{aligned}
&T_{walk} = \frac{2}{5} \\
&T_{scooter} = \frac{2}{10} \\
&E(T) = 0.4 \times \frac{2}{5} + 0.6 \times \frac{2}{10} = 0.28
\end{aligned}
$$

Problem 3

$$
\begin{aligned}
&X = F^{-1}(U) \\
&P(X \le x) = P(F^{-1}(U) \le x) = P\big(U \le F(x)\big) = F(x)
\end{aligned}
$$

because $P(U \le u) = u$, when U is uniform on $[0,1]$

https://en.wikipedia.org/wiki/Inverse_transform_sampling#Proof_of_correctness

Problem 4

Part A

$$\because X_N \sim Bi(N, P)$$
$$E(\hat{p}_N) = E\left(\frac{X_N}{N}\right) = \frac{E(X_N)}{N} = \frac{NP}{N} = P$$
$$Var(\hat{p}_N) = Var\left(\frac{X_N}{N}\right) = \frac{Var(X_N)}{N^2} = \frac{NP(1-P)}{N^2} = \frac{P(1-P)}{N}$$
$$sd(\hat{p}_N) = \sqrt{Var(\hat{p}_N)} = \sqrt{\frac{P(1-P)}{N}}$$

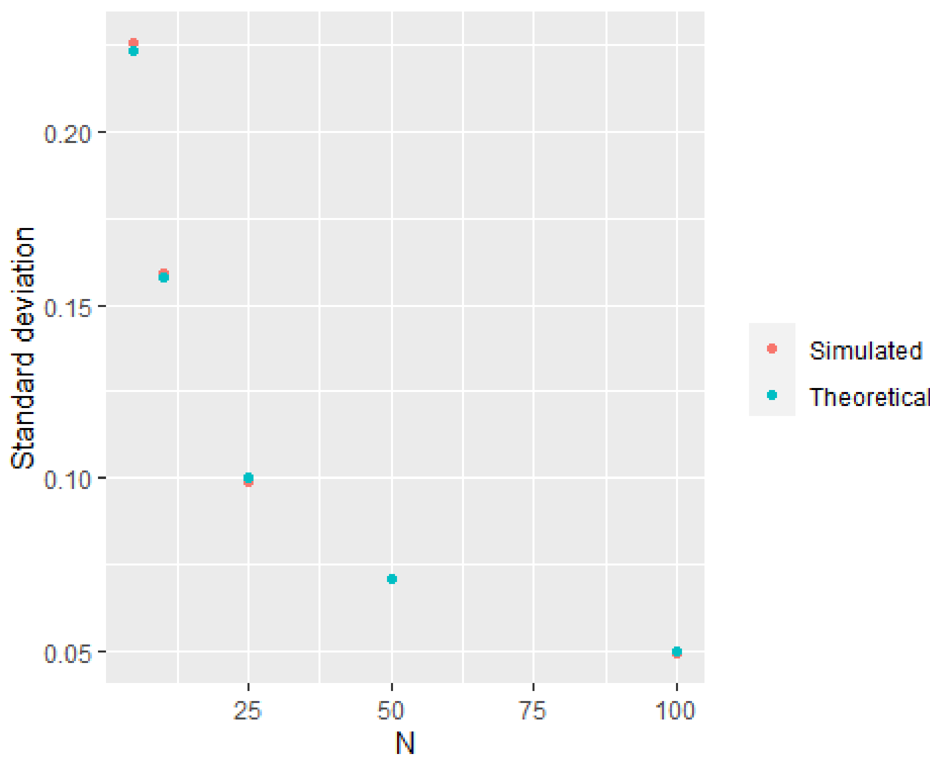
Part B

Theoretical $E(\hat{p}_5) = 0.5, sd(\hat{p}_5) = \sqrt{\frac{0.5\times(1-0.5)}{5}} = 0.2236$

Simulated mean and sd of $\hat{p}_5$

```
## [1] 0.49754

## [1] 0.2257765
```



Comment: Standard deviation decreases as N increases.

Problem 5

$$\because Y = \max\{X_1, \dots, X_N\}$$

$\therefore X_i \leq Y$ for all $i = 1, \dots, N$

$$F(Y) = P(Y \leq x) = P(X_1 \leq x, X_2 \leq x, \dots, X_N \leq x)$$

$= P(X_1 \leq x)P(X_2 \leq x) \cdots P(X_N \leq x)$ (independent)

$= [P(X \leq x)]^N$ (identical)

$$= [F(X)]^N$$

$$\because F(X) = 1 - e^{-\lambda x}$$

$$\therefore F(Y) = [F(X)]^N = \left(1 - e^{-\lambda x}\right)^N$$

$$\therefore f(Y) = \frac{d}{dx}F(Y) = N\left(1 - e^{-\lambda x}\right)^{N-1} \cdot \left(\lambda e^{-\lambda x}\right)$$

https://en.wikipedia.org/wiki/Order_statistic