CS 70 HW #13

Sunday – Cyuo@berkeley.edu

Question I

a) Using the givens:

$$P(X \geq \alpha) \leq \frac{E[(X+c)^2]}{(\alpha+c)^2}$$

$$= \frac{E[X^2] + 2cE[X] + c^2}{(\alpha+c)^2}$$

$$= \frac{(\sigma^2) + (c^2)}{(\alpha+c)^2}$$

b) To solve this, we need to get the derivative and set it to 0 in respect to c

$$\frac{d}{dc}\left(\frac{(\sigma^2) + (c^2)}{(\alpha+c)^2}\right) = 0$$

$$\downarrow$$

$$\frac{2c(\sigma+c)^2 - 2(\alpha+c)(\sigma^2+c^2)}{(\alpha+c)^4} = 0$$

$$\downarrow$$

$$2c(\sigma+c)^2 - 2(\alpha+c)(\sigma^2+c^2) = 0$$

$$\downarrow$$

$$\alpha c^2 + c(\alpha^2 - \sigma^2) - \sigma^2\alpha = 0$$

$$\downarrow$$

$$c = \frac{\sigma^2}{\alpha}$$

c) i) It isn't always correct because there is leeway in the variance we get, or rather, it wouldn't always return 50% odds

ii) We can set RV: X

X: one fulls has higher probability than another

$P(X=0) = 0.75$ $\quad$ $P(X=4) = 0.25$

w/ $\alpha = 2$

↳ $E[X] = 0.25$

↳ variance $= 100(0.25)(0.75)$

$\downarrow$

$P[X \geq E[X] + 7] = \frac{1}{4} > \frac{Var(X)}{2.7^2} = 0.19$

d) i) markov: $P(X \geq 5) \leq \frac{E[X]}{5} = \frac{3}{5}$

ii) chebyshev: $P(X \geq 5) \leq P[|X - E[X]| \geq 2] \leq \frac{VarX}{2^2} = \frac{1}{2}$

iii) Since $E[X] = E[Y]$

↳ $Y = X - E[X] = X - 3$

↳ $P(X \geq 5) = P(Y \geq 2) \leq \frac{VarY}{2^2 + VarY} = \frac{1}{3}$

Cantelli > chebyshev > markov   in order of effectiveness of bounds

Question 2

Need find variance

X_i: number of visits needed to collect the ith monopoly card given $(i-1)$ Monopoly cards already collected

$p = (n-i+1)/n$

$\downarrow$

$$\mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

$$= \sum_{i=1}^{n} \frac{1-(n-i+1)/n}{[(n-i+1)/n]^2}$$

$$= \sum_{j=1}^{n} \frac{1-j/n}{[j/n]^2}$$

$$= \sum_{j=1}^{n} \frac{n(n-j)}{j^2}$$

$$= \sum_{j=1}^{n} \frac{n^2}{j^2} - \sum_{j=1}^{n} \frac{n}{j}$$

$$= \boxed{n^2 \sum_{j=1}^{n} \frac{1}{j^2} - E[X]}$$

Question 3

a) task 1: estimate $2/p_1$

$\hookrightarrow \frac{2}{\pi} = p_1 \rightarrow 2/p_1 = \pi$

task 2: estimate $4p_2$

$\hookrightarrow p_2 = \frac{\pi}{4} \rightarrow 4p_2 = \pi$

task 3: $\sqrt{\frac{6}{p_3}}$ estimate

$p_3 = \frac{6}{\pi^2} \rightarrow \pi^2 = \frac{6}{p_3} \rightarrow \pi = \sqrt{\frac{6}{p_3}}$

b) $P(|X_n - \mu| \geq \varepsilon) \leq \frac{1}{15^2}$; $\frac{1}{15^2} = \delta$

$\hookrightarrow n \geq \frac{\sigma^2}{\varepsilon^2 \delta}$

$\hookrightarrow n \geq \left(\pi \frac{(4-\pi)}{16} \div \varepsilon^2 \delta\right)$

$\hookrightarrow n \geq \left(\frac{\pi(4-\pi)}{16\,\varepsilon^2(\delta)}\right)$

d) $i=1$

↳ $P[|x_i - \pi|] = P[n_1 \sim 2/\pi]$

↳ using Chernoff inequality ↓

$(1-\frac{2}{\pi})(1-\pi)/\epsilon N \leq 2$

↳ $N \geq \frac{2}{\pi}(1-\frac{2}{\pi})/\epsilon_2$

$i=2$

$P[|x_i - \pi|] = P[n_2 \sim \frac{6}{\pi^2}]$

↳ using Chernoff inequality ↓

$(1-\pi^2/2)(1-\pi_2)/\epsilon N \leq 2$

↳ $N \geq \frac{\pi^2}{6}(1-\frac{\pi^2}{6})/\epsilon_2$

e) from this, we can see that technique 1 has the lowest and technique 2 has the highest N values

Question 4

a) Because x is uniform

$\hookrightarrow$ $f_X(x) = \begin{cases} 1/2 & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$ PDF

CDF $F(x) = P[X \leq x] = \begin{cases} 0 ; & x < 0 \\ x/2 ; & 0 \leq x < 2 \\ 1 ; & x \geq 2 \end{cases}$

However looking at y

$\hookrightarrow$ $y = 4x^2 + 1$

(Y) CDF: $P\left(X \leq \frac{\sqrt{y-1}}{2} \right) = F_Y(y)$

$\hookrightarrow$ $F_Y(Y) = \begin{cases} 0 ; & y < 1 \\ \frac{\sqrt{y-1}}{2} ; & 1 \leq y < 17 \\ 1 ; & y \geq 17 \end{cases}$

PDF (Y):

$f_Y(y) = \frac{d}{dy}\left(F_Y(y) \right)$

$= \begin{cases} \frac{1}{4\sqrt{y-1}} ; & 1 \leq y \leq 17 \\ 0 ; & \text{otherwise} \end{cases}$

Variance / Example :

$$E(x) = \frac{0+2}{2} = 1$$

$$Var(x) = \frac{1}{12}(2-0)^2 \rightarrow \frac{1}{3}$$

For Y :

$$E(Y) = E(4x^2+1) \rightarrow 16E(x) + E(1)$$

$$\hookrightarrow 16(1) + 1$$

$$\hookrightarrow 17$$

$$Var(Y) = Var(4x^2+1) = 18\ Var(x)$$

$$\hookrightarrow 6 = Var(Y)$$

b)
$$f_{x,y}(x,y) = \begin{cases} cxy + 1/4 & ; x \in \{1,2\} \\ 0 & ; y \in [0,2] \\ & ; \text{otherwise} \end{cases}$$

$$\int_x \int_y f_{x,y}(x,y)\, dy\, dx = 1$$

$$\hookrightarrow \int_1^2 \left(2cx + \frac{1}{2}\right) dx = 1$$

$$\hookrightarrow 2c\left(\frac{x^2}{2}\right)\Big|_1^2 + \left(\frac{x}{2}\right)\Big|_1^2 = 1$$

$$\hookrightarrow 2c\left(2 - \frac{1}{2}\right) + \left(1 - \frac{1}{2}\right) = 1$$

$$\hookrightarrow 3c = 1/2$$

$$\hookrightarrow c = 1/6$$

↳ X and Y are Independent because we used the rule shown [illegible] in [illegible]

c) $X \sim Exp(3)$

which mean $\lambda = 3$

i) $Y = \lceil X \rceil \rightarrow Geom(p)$ where $p = 1 - e^{-3}$

ii) Thus $p_X(Y=k) = (1-p)^k p$

↳ when $p = 1 - e^{-3}$ so in this case, a discrete distribution

and it is geometric distribution

d) $X_i \sim Exp(\lambda_i)$ for i mutually independent

$$P[(\min(X_1 \ldots X_n) > x]$$

$$\prod_{i=1}^{n} \exp(-x \lambda_i)$$

↳ $\exp(-x \sum_{i=1}^{n} \lambda_i)$

$$\min(X_1 \ldots X_n) \sim Exp\left(\sum_{i=1}^{n} \lambda_i\right)$$

Question 5

a) $X: P[x \in (a,b)] > 0 \; ; a,b \in \mathbb{R}$

↳ $X \sim \mathcal{N}(0,1)$ would be an example that fits the model

↳ $P[x \in (a,b)] = F[b] - F[a]$ for $\forall b > a$

Since this is an increasing function, the layout of will inevitably fall under the given conditions

b) The function is increasing due to the following:

$f(x+c) > f(x) \; ; \text{ for any } c$

$= P[X < x+c] - P[X < x]$

$= P[X < x] + P[X \in (x, x+c)] - P[X < x]$

$= P[X \in (x, x+c)] > 0$

↳ From this, we can see that the function is strictly increasing

c) Unif $(0,1)$

$F^{-1}(U) = Y$

↳ $F_Y(y) = P[Y \leq y]$

$= P[F^{-1}(U) \leq y]$

$= P[F[F^{-1}(U)] \leq F(y)]$

$= P[U \leq F(y)]$

$= F(y) = P[X \leq y]$

$= F_X(y)$

↳ From this, we can see that the variant comes out to be true

2) X is a discrete random variable

w/ $P[X = x_i] = p_i$

i) generate unif $[0,1)$

ii) $Y = \begin{cases} x_1, & \text{if } 0 \le U < p_1 \\ x_2 & \text{if } p_1 \le U < p_1 + p_2 \\ \vdots & \\ x_i & \text{if } p_1 + \dots p_i \le U < p_1 + p_2 + \dots p_i \\ x_n & \text{if } p_1 + \dots p_{n-1} \le U < p_i + \end{cases}$

↳ $Y = x_i$ w/ probability $= P[p_1 + \dots p_i]$

↳ $Y \stackrel{d}{=} X$

→ By this method, u can generate X

Question 6

a) $$P[Y_i < X_i] = \int_{t=0}^{\infty} P[X_i = t \cap Y_i < t]$$

$$= \int_{t=0}^{\infty} P[X_i = t] P[Y_i < t]$$

$$= \int_{t=0}^{\infty} \lambda e^{-\lambda t} (1 - e^{-\mu t})$$

$$= \int_{t=0}^{\infty} \lambda e^{-\lambda t} - \lambda e^{-(\lambda + \mu) t}$$

$$= \int_{t=0}^{\infty} \lambda e^{-\lambda t} - \int_{t=0}^{\infty} \lambda e^{-(\lambda + \mu) t}$$

$$= 1 - \frac{\lambda}{\lambda + \mu}$$

$$= \frac{\mu}{\mu + \lambda}$$

b) $$P(D > 2) = P(X > 20 + 2 \mid X \geq 20)$$

$\downarrow$ Bayes Rule

$$P(X > 20 + 2 \mid X \geq 20) = \frac{P(X > 20 + 2)}{P(X \geq 20)}$$

$$= \frac{e^{\mu(20+2)}}{e^{20\mu}}$$

$$= e^{\mu 2}$$

↳ CDF Distribution : $1 - e^{\mu t}$

* D is exponentially distributed w/ parameter μ

c) whichever ordering first is what Z will be, so:

$Z = \min(X, Y)$

$$P(Z > T) = P(X > T \cap Y > T)$$
$$= P(X > T)\, P(Y > T)$$
$$= (1 - P(X \leq t))(1 - P(Y \leq t))$$
$$= (1 - (1 - e^{-\mu T}))(1 - (1 - e^{-\lambda t}))$$
$$= e^{-\mu t}\, e^{-\lambda t}$$
$$= e^{-(\mu + \lambda)t}$$

↳ CDF follows for Z : $P(Z \leq t) = 1 - e^{-(\mu + \lambda)t}$

↳ Z exponential w/ parameters μ and λ

d) if $t > 0$

↳ if $x_1 + x_2 \leq t \rightarrow x_1, x_2 \geq 0 \rightarrow$ $x_1 \leq t$, $x_2 \leq t - x_1$

↓ Subsequently

$P(T < t) =$

$$P(x_1 \leq t, x_2 \leq t - x_1) = \int_0^t \int_0^{t-x_1} \lambda \exp(-\lambda x_1)(\lambda \exp(-\lambda x_2)) \, dx_2 \, dx_1$$

$$= \lambda^2 \int_0^t \exp(-\lambda x_1) \cdot \frac{1 - \exp(-\lambda(t - x_1))}{\lambda} \, dx_1$$

$$= \lambda \int_0^t \exp(-\lambda x_1) - \exp(-\lambda t) \, dx_1 = \lambda \left[\frac{1 - \exp(-\lambda t)}{\lambda} - t \exp(-\lambda t) \right]$$

Differentiating CDF

$$f_U(t) = \frac{d}{dt} P(T < t) = \lambda \exp(-\lambda t) - \lambda \exp(-\lambda t) + \lambda^2 t \exp(-\lambda t)$$

$$= \lambda^2 t \exp(-\lambda t)$$