CS 70 Bounds Review

Markov Inequality:

$$P[X \geq k] \leq \frac{\mathbb{E}[X]}{k}$$

only holds for non-negative RVs X

Chebyshev's Inequality:

$$P[|X-\mu| \geq \varepsilon] \leq \frac{Var(X)}{\varepsilon^2}$$

Alternately,

$$P[|X-\mu| \geq k\sigma] \leq \frac{1}{k^2}$$

$\boxed{1}$ X: RV number of balls that land in first bin

(a)

$X \sim \text{Binomial}(n, p)$

n: number of balls

$p = \frac{1}{n}$

$$\mathbb{E}[X] = np = n\left(\frac{1}{n}\right) = \boxed{1}$$

(b)

$$P[X \geq k] \leq \frac{\overbrace{\mathbb{E}[X]}^{1}}{k} \qquad \boxed{\frac{1}{k}}$$

(c) $$Var(X) = np(1-p) = n\left(\frac{1}{n}\right)\left(1-\frac{1}{n}\right) = \boxed{\frac{n-1}{n}}$$

(d) $P[X \geq k]$

$$P[|X-\mu| \geq \varepsilon] \leq \frac{Var(X)}{\varepsilon^2}$$

$$P[|X-1| \geq k-1] \leq \frac{Var(X)}{(k-1)^2} = \boxed{\frac{n-1}{n(k-1)^2}}$$

[2] m balls $\to$ n bins   distinguishable

(a) $P[\text{at least 1 collision in bin 1}]$

$= 1 - P(0 \text{ balls in bin } 1) - P(1 \text{ ball in bin } 1)$

$$= \boxed{1 - \left(\frac{n-1}{n}\right)^m - \binom{m}{1}\left(\frac{1}{n}\right)\left(\frac{n-1}{n}\right)^{m-1}}$$

(b) N = RV for number of bins with collisions

$P[N \geq \frac{n}{2}]$

Let $X_1, X_2 \ldots X_n$ be indicator RVs for whether or not a particular bin has a collision

$N = X_1 + \ldots + X_n$

$$X_i = \begin{cases} 1 & \text{wp } p \\ 0 & \text{wp } 1-p \end{cases}$$

wp = with probability

$\mathbb{E}[N] = \mathbb{E}[X_i] = np$

$\mathbb{E}[X_i] = 1(p) = p$

$P[N \geq \frac{n}{2}] \leq \frac{\mathbb{E}[N]}{n/2}$

$= \frac{np}{n/2} = \boxed{2p}$

(c) K_i = number of ~~keys~~ balls assigned to bin i

$\text{Var}(K_i)$

$$K_i \sim \text{Binom}(\overset{\text{num trials}}{m}, \underset{\text{prob. of success}}{\frac{1}{n}})$$

$$np(1-p)$$

$$= m\left(\frac{1}{n}\right)\left(\frac{n-1}{n}\right) = \boxed{\frac{m(n-1)}{n^2}}$$

(d) $P\left[K_i \geq \frac{3m}{n}\right]$

$$P[|X-\mu| \geq \varepsilon] \leq \frac{\text{Var}(X)}{\varepsilon^2}$$

$$\mathbb{E}[K_i] = np = m\left(\frac{1}{n}\right) = \frac{m}{n}$$

$$= P\left[\left|K_i - \frac{m}{n}\right| \geq \frac{2m}{n}\right]$$

$$\leq \frac{\text{Var}(K_i) \leftarrow v}{\left(\frac{2m}{n}\right)^2} = \boxed{\frac{n^2 v}{4m^2}}$$

$$\left|K_i - \frac{m}{n}\right| \geq \frac{2m}{n}$$

$$K_i - \frac{m}{n} \geq \frac{2m}{n}$$

$$K_i \geq \frac{3m}{n}$$

$\boxed{3}$ $A = \frac{X_1 + \dots + X_n}{n}$, where X_i denotes the number of ppl entering the store in the i^{th} hour

$$\mathbb{E}[A] = \mathbb{E}\left[\frac{X_1 + \dots + X_n}{n}\right] = \frac{1}{n}\mathbb{E}[X_1 + \dots + X_n] = \frac{1}{\cancel{n}}(\cancel{n})\overset{\lambda}{\cancel{\mathbb{E}[X_i]}}$$

$$= \underline{\lambda}$$

$$\text{Var}(A) = \text{Var}\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n^2}\text{Var}(X_1 + \dots + X_n) = \frac{1}{n^2}n\left(\overset{\lambda}{\cancel{\text{Var}\,X_i}}\right) = \frac{\lambda}{n}$$

$$P(|\underline{A} - \underset{\lambda}{\mu}| \geq \underset{\varepsilon = 1}{\varepsilon}) \leq \frac{Var(A)}{\varepsilon^2} \leq 0.05$$

$\frac{\lambda}{n}$

width 2

$$P(|A - \lambda| \geq 1) \leq \frac{\lambda}{n} \leq 0.05$$

$[a, b]$ $b - a = 2$

$$\lambda \leq 0.05n$$

$$n \geq 20\lambda \quad (\lambda \text{ is at most } 10)$$

$$\boxed{n \geq 200} \text{ hours}$$

$\boxed{4}$ $\lambda \geq 10$ Once again our sample average RV is:

$$A = \frac{X_1 + \ldots + X_n}{n}$$

$$\mathbb{E}[A_n] = \frac{1}{n}\mathbb{E}[X_1 + \ldots + X_n] = \frac{1}{\cancel{n}}(\cancel{n})\frac{1}{\lambda} = \frac{1}{\lambda}$$

$$Var(A_n) = \frac{1}{n^2}Var(X_1 + \ldots + X_n) = \frac{1}{n^2}n\left(\frac{1}{\lambda^2}\right) = \frac{1}{n\lambda^2}$$

$$P(|A_n - \frac{1}{\lambda}| \geq \varepsilon^2) \leq \frac{Var(A_n)}{\varepsilon^2} \leq 0.05 \quad (\text{where } \varepsilon \text{ is unknown})$$

$$\frac{1}{n\lambda^2\varepsilon^2} \leq \frac{1}{20}$$

$$n\lambda^2 e^2 \geq 20$$

$$\varepsilon^2 \geq \frac{20}{n\lambda^2}$$

$$\varepsilon \geq \frac{2\sqrt{5}}{\lambda\sqrt{n}} \longleftarrow (2\sqrt{5} \approx 4.5)$$

Our confidence interval: $(A - \varepsilon, A + \varepsilon)$

$$= \left(A - \frac{4.5}{\lambda\sqrt{n}},\ A + \frac{4.5}{\lambda\sqrt{n}}\right)$$

However since $\lambda \geq 10$, $= \boxed{\left(A - \frac{.45}{\sqrt{n}},\ A + \frac{.45}{\sqrt{n}}\right)}$