Chebyshev's Inequality: Let X be a r.v. with mean μ and variance σ^2. Then for any $k > 0$

$$P(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

or

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

__Proof__ : Let X be continuous with pdf $f_X(x)$. $E(X) = \mu$,

$$\sigma^2 = Var(X) = E(X-\mu)^2$$

$$= \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\, dx$$

$$= \int_{|x-\mu| \geq k} \cdots + \int_{|x-\mu| < k} \cdots$$

$$\geq \int_{|x-\mu| \geq k} (x-\mu)^2 f_X(x)\,dx \geq k^2 \int_{|x-\mu| \geq k} f_X(x)\,dx$$

$$= k^2 \, P(|X-\mu| \geq k)$$

So

$$P(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

Other versions

$$P(|X-\mu| \leq k\sigma) \geq 1 - \frac{1}{k^2} \quad (*)$$

or $$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

In $(*)$ let us take $k = 3$

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \geq 1 - \frac{1}{9} = \frac{8}{9}$$

$$\approx 0.89$$

If we take $k=2$, then

$$P(\mu - 2\sigma \le X \le \mu + 2\sigma) \ge 1 - \frac{1}{4} = \frac{3}{4}$$

$$= 0.75$$

Examples: The number of customers who visit a store (per hour) is a r.v. X with $\mu = 18$ and $\sigma = 2.5$. With what prob. can we assert that in a given hour between 8 to 28 customers will visit?

$k = 4$

By Chebyshev's inequality (*), take

$k=4$

$$P(8 \leq X \leq 18) = P(|X-18| \leq 4 \times 2.5)$$

$$\geq 1 - \frac{1}{16} = \frac{15}{16} = 0.9375$$

2. Suppose independent observations are taken from a population with mean μ and variance 1. How many observations are needed in order that

prob. is at least 0.9 that the mean of observations differs from μ by not more than 1 ?

<u>Sol^n</u> Let $X_1, \ldots, X_n$ be n observations

$$E(X_i) = \mu, \quad V(X_i) = 1$$

We want

$$P(|\overline{X} - \mu| < 1) \geq 0.9 \qquad ??$$

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$E(\overline{X}) = \frac{1}{n} E(X_1 + \cdots + X_n)$$

$$= \frac{1}{n}\left(\mu + \cdots + \mu\right) = \mu$$

$$V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \quad 1$$

$$Var(X+Y) = E\left(X+Y - E(X+Y)\right)^2$$

$$= E\left[\{X - E(X)\} + \{Y - E(Y)\}\right]^2$$

$$= E(X - E(X))^2 + E((Y - E(Y))^2$$

$$+ 2E\{(X - E(X))(Y - E(Y))\}$$

$$= Var(X) + Var(Y) + \underbrace{2\,Cov(X,Y)}$$

For independent r.v. $Cov(X,Y) = 0$

So for independent r.v's $X_1 \ldots, X_n$

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

$$Var(\bar{X}) = Var\left(\frac{1}{n}\sum X_i\right)$$

$$= \frac{1}{n^2} V\left(\sum X_i\right) = \frac{1}{n^2}\sum V(X_i)$$

$$= \frac{1}{n^2}\cdot(\sigma^2 + \cdots + \sigma^2) = \frac{\sigma^2}{n}$$

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

$$P(|\bar{X} - \mu| < 1) \geq 1 - \frac{1}{n}$$

We want $1 - \frac{1}{n} > 0.9$

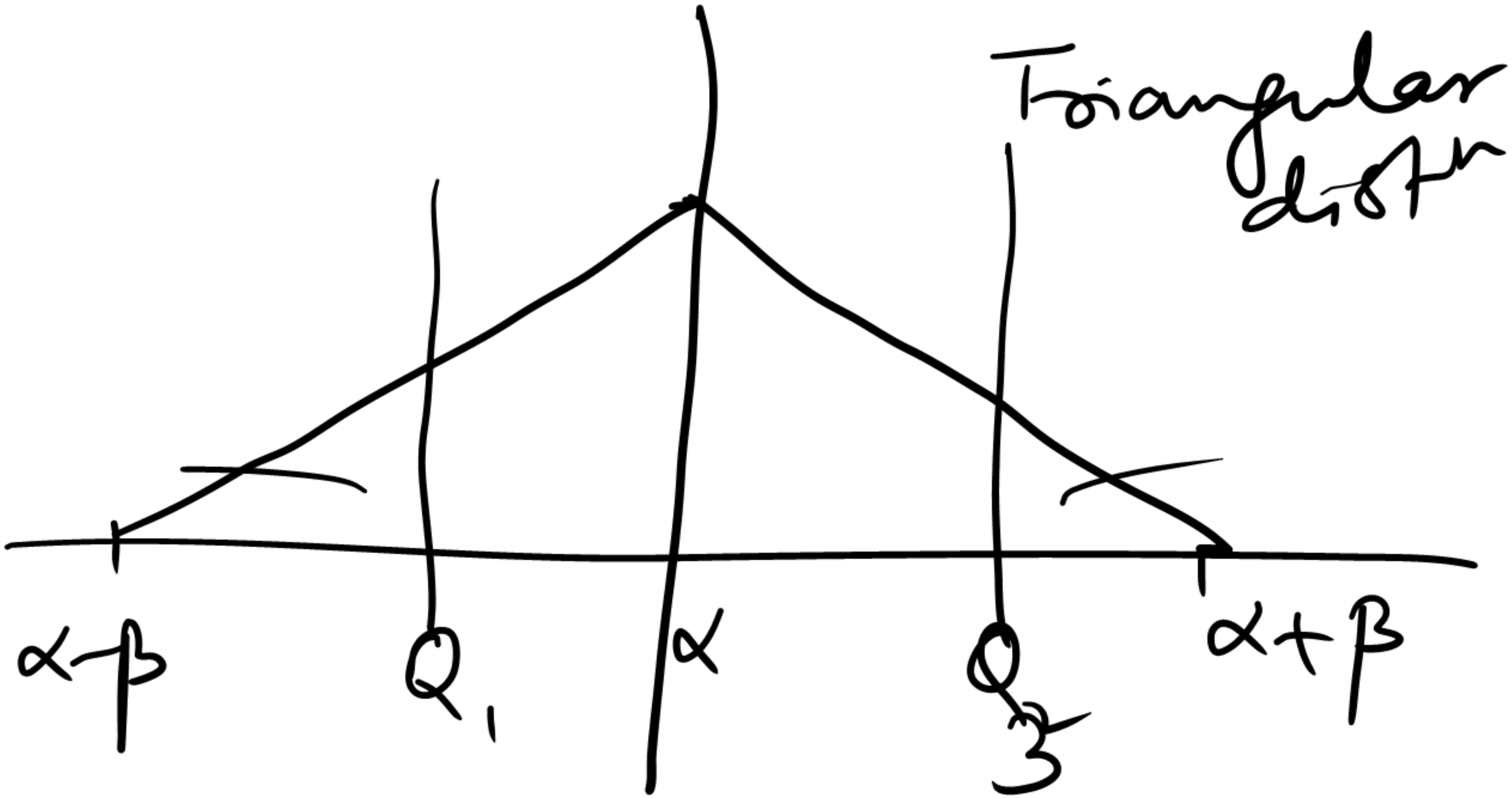
$$\Rightarrow n > 10$$

<u>Example</u>: Moments, median, quartiles

Let X be a continuous r.v.

with

$$f_X(x) = \frac{1}{\beta}\left\{1 - \frac{|x-\alpha|}{\beta}\right\}, \quad \alpha-\beta < x < \alpha+\beta$$



$$Med(X)=\alpha, \qquad E(X)=\alpha$$

$$Var(X)= E(X-\alpha)^2$$

$$= \int_{\alpha-\beta}^{\alpha+\beta} (x-\alpha)^2 \frac{1}{\beta} \left\{ 1 - \frac{|x-\alpha|}{\beta} \right\} dx$$

$$y = \frac{x-\alpha}{\beta}$$

$$dy = \frac{1}{\beta} dx$$

$$= \beta^2 \int_{-1}^{1} y^2 (1-|y|) \, dy$$

$$= 2\beta^2 \int_0^1 y^2 (1-y) \, dy = \frac{\beta^2}{6}$$

$\underline{\text{cdf}}$: $F_X(x) = \displaystyle\int_{-\infty}^{x} f_X(t)\,dt$

$= 0$ if $x \le \alpha - \beta$

$\alpha - \beta < x < \alpha + \beta$

$$F_X(x) = \int_{\alpha-\beta}^{x} \frac{1}{\beta}\left\{1 - \frac{|t-\alpha|}{\beta}\right\} dt$$

$$y = \frac{t-\alpha}{\beta}$$

$$= \int_{-1}^{\frac{x-\alpha}{\beta}} (1 - |y|)\, dy$$

If $\quad x < \alpha$, then

$$I = \int_{-1}^{(x-\alpha)/\beta} (1+y)\, dy = \left.\frac{(1+y)^2}{2}\right|_{-1}^{\frac{x-\alpha}{\beta}}$$

$$= \frac{1}{2}\left[1 + \left(\frac{x-\alpha}{\beta}\right)\right]^2, \quad \alpha-\beta < x < \alpha$$

$x = \alpha, \quad I = \frac{1}{2}$

$$\alpha \leq x < \alpha+\beta$$

$$I = \frac{1}{2} + \int_{\alpha}^{x} \frac{1}{\beta}\left(1 - \frac{|t-\alpha|}{\beta}\right)dt$$

$$= \frac{1}{2} + \int_0^{(x-\alpha)/\beta} (1-y)\,dy$$

$$= 1 - \frac{1}{2}\left(1 - \frac{x-\alpha}{\beta}\right)^2$$

So

$$F_X(x) = 0, \quad x \le \alpha - \beta$$

$$= \frac{1}{2}\left[1 + \frac{x-\alpha}{\beta}\right]^2, \quad \alpha-\beta < x \le \alpha$$

$$= 1 - \frac{1}{2}\left[1 - \frac{x-\alpha}{\beta}\right]^2, \quad \alpha < x < \alpha+\beta$$

$$= 1, \quad x \geq \alpha + \beta$$

Suppose we want to find quartiles Q_1 and Q_3.

$$F_X(Q_1) = \frac{1}{4}$$

$$\frac{1}{2}\left[1 + \frac{Q_1 - \alpha}{\beta}\right]^2 = \frac{1}{4}$$

$$\Rightarrow \quad Q_1 = \alpha + \beta\left(1 - \frac{1}{\sqrt{2}}\right)$$

$\ast$ Ex : Check these

$$F_X(Q_3) = \frac{3}{4}$$

$$1 - \frac{1}{2}\left[1 - \frac{x-\alpha}{\beta}\right]^2 = \frac{3}{4}$$

Ⓡ Find measures of skewness & kurtosis

<u>Problems</u> 1. A student appears in

Mathematics & Statistics tests on the same day. The prob. of passing at least one of the tests is 4/5, whereas the prob of passing both the tests is $\frac{14}{45}$.

If the prob of passing Maths test is $2/3$, find the prob of passing the Stat. test

Sol^n. $P(M \cup S) = \frac{4}{5}$,

$P(M \cap S) = \frac{14}{45}$, $P(M) = \frac{2}{3}$

$$P(S) = P(M \cup S) - P(M) + P(M \cap S)$$

$$= \frac{4}{9}.$$

2. Ramesh and Pooja throw alternately a pair of dice. Ramesh wins if he throws a sum of 6 before Pooja throws a sum of 7. Pooja wins if she throws a sum of 7 before Ramesh throws a sum of 6. Find the prob of

Ramesh winning the game if he starts the game. Similarly for Pooja if she starts the game.

$$P(\text{throwing a } 6) = \frac{5}{36}$$

$$P(\text{throwing a } 7) = \frac{6}{36} = \frac{1}{6}$$

$$P(\text{Ramesh wins}) = P(\text{wins on I}^{st}) + P(\text{wins on III}^{rd}) + \cdots$$

$$= \frac{5}{36} + \frac{31}{36} \times \frac{5}{6} \times \frac{5}{36}$$

$$+ \left(\frac{31}{36} \times \frac{5}{6}\right)^2 \times \frac{5}{36} + \cdots$$

$$= \frac{5}{36} \cdot \frac{1}{1 - \frac{31}{36} \times \frac{5}{6}} = \frac{30}{61}$$

$$P(\text{Pooja wins if she starts}) = \frac{36}{61}$$

3. Let $A_1, \ldots, A_{10}$ be mutually exclusive and exhaustive events with $P(A_k)$ proportional to k, $k = 1, \ldots, 10$. Find the conditional probability that at least one of $A_2, A_4, A_6, A_8, A_{10}$ occurs given that at least one of $A_1, A_2, \ldots, A_5$ have occurred.

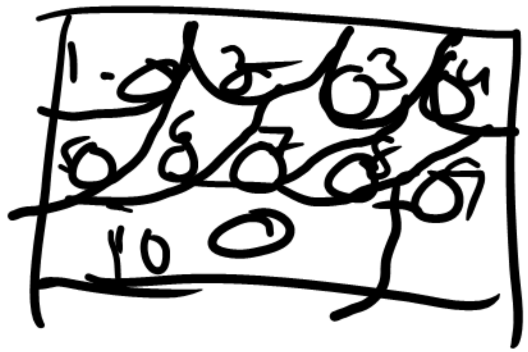
<u>Sol^n</u> $P(A_i) = \alpha i, \quad i = 1 \ldots 10$

$$\alpha \sum_{1}^{10} i = 55\alpha = 1 \Rightarrow \alpha = \frac{1}{55}$$

$$P\left(A_2 \cup A_4 \cup A_6 \cup A_8 \cup A_{10} \,\middle|\, \bigcup_{i=1}^{5} A_i\right)$$

$$= \frac{P\left(\left(\bigcup_{i=1}^{5} A_{2i}\right) \cap \left(\bigcup_{i=1}^{5} A_i\right)\right)}{P\left(\bigcup_{i=1}^{5} A_i\right)}$$

$$= \frac{P(A_2 \cup A_4)}{\sum_{i=1}^{5} P(A_i)}$$



$$= \frac{P(A_2) + P(A_4)}{\frac{1}{55}\sum_{i=1}^{5} i} = \frac{\frac{2}{55} + \frac{4}{55}}{\frac{15}{55}}$$

$$= \frac{6}{15} = \frac{2}{5}.$$

4. Let X be discrete r.v. with p.m.f

$$p_X(k) = \alpha \frac{e^{-1}}{k!}, \quad k = 1, 2, \ldots$$

Find α and find $E\left(\frac{1}{1+X}\right)$

$$\alpha \sum_{k=1}^{\infty} \frac{e^{-1}}{k!} = 1$$

$$\Rightarrow \alpha e^{-1} \left(\frac{1}{1!} + \frac{1}{2!} + \cdots \right) = 1$$

$$\Rightarrow \alpha e^{-1} (e - 1) = 1$$

$$\Rightarrow \alpha = \frac{1}{1 - e^{-1}}$$

$$E\left(\frac{1}{1+X}\right) = \sum_{k=1}^{\infty} \frac{1}{(1+k)} \cdot \frac{1}{(1-e^{-1})} \cdot \frac{e^{-1}}{k!}$$

$$= \frac{e^{-1}}{1-e^{-1}} \left[\sum_{k=1}^{\infty} \frac{1}{(k+1)!} \right]$$

$$= \left(\frac{e^{-1}}{1-e^{-1}} \right) \left[\frac{1}{2!} + \frac{1}{3!} + \cdots \right]$$

$$= \left(\frac{e^{-1}}{1-e^{-1}} \right) \left(e - 2 \right) = \frac{e-2}{e-1}$$

$\subset$ → 'is subset of'

Here it can be equal also.

$$F(x) = 0, \quad x < 0$$
$$= \frac{x}{4}, \quad 0 \le x < 1$$
$$= \frac{x+1}{4}, \quad 1 \le x < 2$$
$$= \frac{11}{12}, \quad 2 \le x < 3$$

$= 1, \quad x \geq 3$

X is continuous in $(0,1)$ & $(1,2)$ with pdf $f(x) = \frac{1}{4}, \quad \begin{matrix} 0 < x < 1 \\ 1 < x < 2 \end{matrix}$

$P(X=1) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$

$P(X=2) = \frac{11}{12} - \frac{3}{4} = \frac{1}{6}$

$P(X=3) = 1 - \frac{11}{12} = \frac{1}{12}$

$$P\left(\frac{1}{2} < X < \frac{5}{2}\right) = F\left(\frac{5}{2}-\right) - F\left(\frac{1}{2}\right)$$

$$= \frac{11}{12} - \frac{1}{8} = \frac{19}{24}$$

$$P(1 < X < 3) = P(X < 3) - P(X \leq 1)$$

$$= F(3-) - F(1)$$

$$= \frac{11}{12} - \frac{1}{2} = \frac{5}{12}$$

$$E(X) = \int_0^1 \frac{x}{4}\,dx + \int_1^2 \frac{x}{4}\,dx$$

$$+ \frac{1}{4} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{12} \cdot 3$$

$E(X^2) = ??$

For median M:

$$P(X \leq M) \geq \frac{1}{2} \quad (1)$$

$$P(X \geq M) \geq \frac{1}{2} \quad (2)$$

$$(1) \Rightarrow F(M) \geq \frac{1}{2} \Rightarrow M \geq 1$$

② $1 - F(M-) \geqslant \frac{1}{2}$

$\Rightarrow \quad F(M-) \leqslant \frac{1}{2}$

$\Rightarrow \quad M \leq 1$

So $M = 1$

So median $(X) = 1$.