Announcement:

① Group Assignment tomorrow.
↳ will be posted this afternoon
② Exam 1 next wednesday
↳ see post in Campuswire.
↳ TAs will hold review / problem sessions.

Poisson Distribution

- Very common distribution.
- # of occurences of rare events occuring during a particular time window.

Ex: - Number of emails recieved in a given day.
- Number of website visits in given week.
- Number of car accidents on a certain stretch of highway in a given year.

Def Parameter: λ - rate averages occurences / time window.

$X \sim \text{Poisson}(\lambda)$ = # of occurences in a given time window

$R_X = \{0, 1, 3, \ldots\}$

PMF: $P_X(k) = \dfrac{e^{-\lambda} \lambda^k}{k!}$, $k = 0, 1, 2, \ldots$

Does it sum to 1?

$$\sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

Taylor series

Interpretation:

$\Delta = T/n$ n pieces.

$X_i \sim \text{Bernoulli}(p = \lambda/n)$

rare if n is big

λ/n - average number of occurances in $[0, T/n]$.

$$X^{(n)} = X_1 + X_2 + \ldots + X_n \sim \text{Binomial}(n, \lambda/n).$$

total # of occurances

$$P_{X^{(n)}}(k) = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\lambda/n\right)^{n-k}$$

$$= \frac{n!}{(n-k)!\,k!}\,\frac{\lambda^k\,(1-\lambda/n)^n}{n^k\,(1-\lambda/n)^k}$$

$$= \frac{n!}{(n-k)!\,n^k} \times \frac{1}{(1-\lambda/n)^k} \times \frac{\lambda^k(1-\lambda/n)^n}{k!}$$

$$\lim_{n\to\infty} \frac{n!}{(n-k)!\,n^k} = 1 \qquad \frac{n!}{(n-k)!\,n^k} = \underset{\substack{\downarrow\\1}}{\left(\frac{n}{n}\right)}\underset{\substack{\downarrow\\1}}{\left(\frac{n-1}{n}\right)} \times \cdots \times \underset{\substack{\downarrow\\1}}{\left(\frac{n-k+1}{n}\right)}$$

$$\lim_{n\to\infty} \left(1-\underset{\substack{\|\\0}}{\lambda/n}\right)^k = 1$$

$$\lim_{n\to\infty} \left(1-\lambda/n\right)^n = e^{-\lambda}$$

$$\lim_{n\to\infty} P_{X^{(n)}}(k) = \frac{e^{-\lambda}\lambda^k}{k!} \quad //$$

__Ex__ Suppose the # of emails I get on a week day is Poisson with rate

$$\lambda = .2 \text{ emails/minute}$$

What is the prob that I get no emails in 5 mins?

Change rate $\lambda = .2(5)$ emails/5 minutes
$= 1$

$$P_X(0) = P(X=0) = \frac{e^{-1} 1^0}{\underbrace{0!}_{=1}} = e^{-1} \approx .3679.$$

Commulation Distribution.

Def: Let X be any R.V. then

$$F_X(x) = P(X \leq x), \quad x \in \mathbb{R}.$$

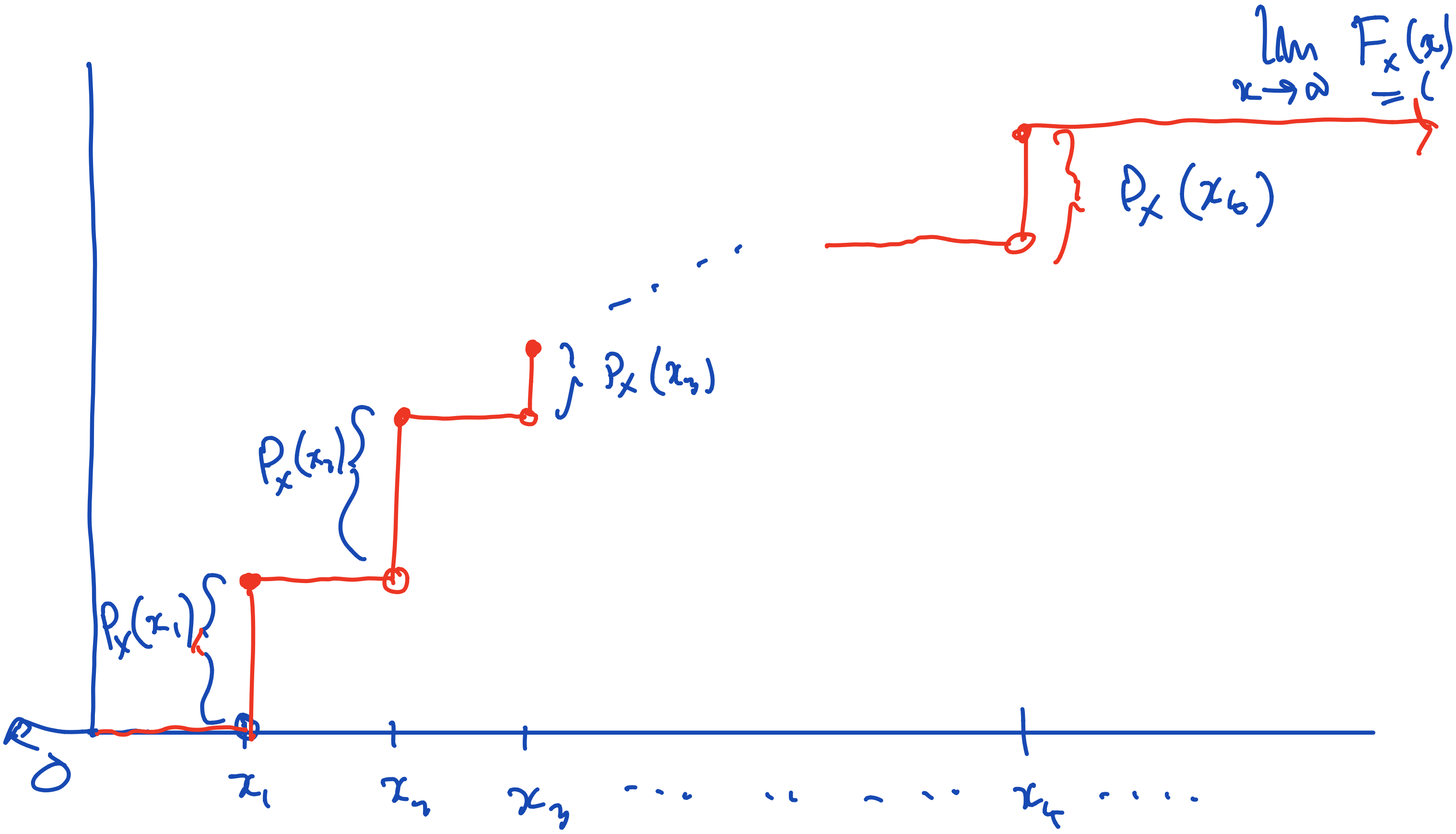
is called the commulation distribution function or CDF of X.

If X is discrete R.V. with

$$R_X = \{x_1, x_2, x_3, \dots\}$$

$x_1 < x_2 < x_3 < \dots$

The CDF of X, $F_X(x)$, looks like



Note

$$F_X(x) = \sum_{\{x_k \le x\}} P_X(x_k).$$

We can get the PMF by.

$$P_X(x_k) = \underbrace{F_X(x_k) - F_X(x_{k-1})}_{\text{jump size}}.$$

Use the CDF to calculate.

$$P(a < X \le b) = P(X \le b) - P(X \le a)$$
$$= F_X(b) - F_X(a).$$

<u>Expected Value.</u>

- What is the average value of a discrete R.V. ?

<u>Def</u> Let X be a discrete R.V. with range. $R_X = \{x_1, x_2, x_3, \dots\}$.

then

$$\mathbb{E}X = \mu_X = \sum_{x_k \in R_X} x_k P(X = x_k).$$
$$= \sum_{x_k \in R_X} x_k P_X(x_k).$$

<u>Interpretation</u> Repeat this experiment N times.

N_k = # of times x_k occurs.

$$P_X(x_k) \simeq \frac{N_k}{N}$$ frequentist interpretation.

$$\text{Average} = \frac{N_1 x_1 + N_2 x_2 + N_3 x_3 + \cdots}{N}$$

$$= x_1 \frac{N_1}{N} + x_2 \frac{N_2}{N} + x_3 \frac{N_3}{N} + \cdots$$

$$\simeq \sum_{k=1}^{\infty} x_k P_X(x_k).$$