Lectures of May 19th, 2006 - PM
Scribe: Saba Siadat Mousavi, Ali Khanafer

1 The Probability Mass Function (PMF) of a Random Variable

The probability mass function (PMF) of a discrete random variable X with range $S_X = \{x_1, x_2, x_3, ...\}$ is defined as:

$$P_X = P[X = x], \text{ for any } x \in S_X$$

Given PMF P_X on S_x, the PDF f_X can be expressed as:

$$f_X = \sum_{x_k \in S_X} P_X(x_k)\delta(x - x_k)$$

1.1 Conditional CDF and PDF

Suppose X is a random variable with CDF F_X and let A be an event. The conditional CDF $F_{X|A}(x)$ of X given A is defined as:

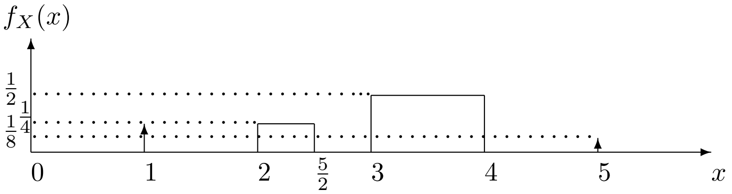
$$F_{X|A}(x) = \frac{P[X \leq x \ and \ A]}{P[A]}$$

The conditional PDF $f_{X|A}(x)$ of X given A is defined as:

$$f_{X|A} \ (x) = \frac{dF_{X|A}(x)}{dx}$$

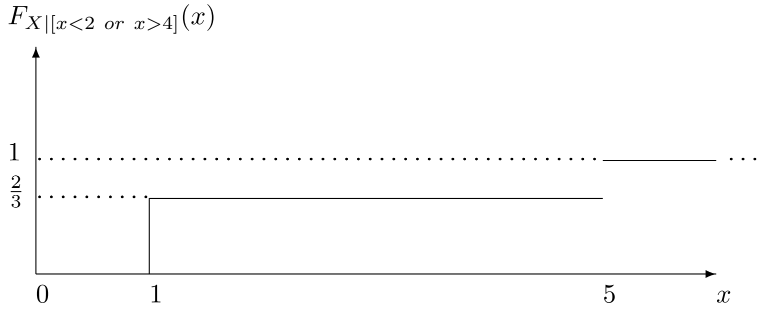
Example:

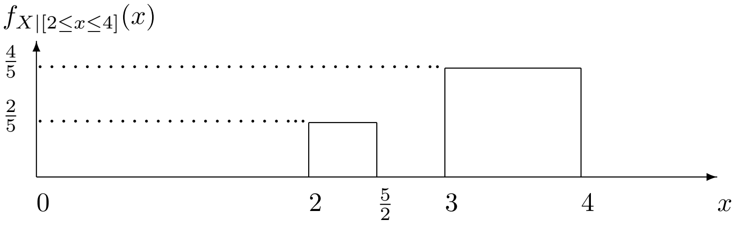
Let f_X of random variable X be given as follows:



1. Find $f_{X|[x<2 \ or \ x>4]}(x)$
2. Find$f_{X|[x\leq 4 \ and \ x\geq 2]}(x)$
3. Find $F_{X|[x<2 \ or \ x>4]}(x)$

Solution:
We mask out the region not in $[x < 2 \ or \ x > 4]$ (i.e. set f_X to zero for $x \in [2,4]$), and then we scale the remaining function so that it integrates to one, in order to qualify the PDF condition.



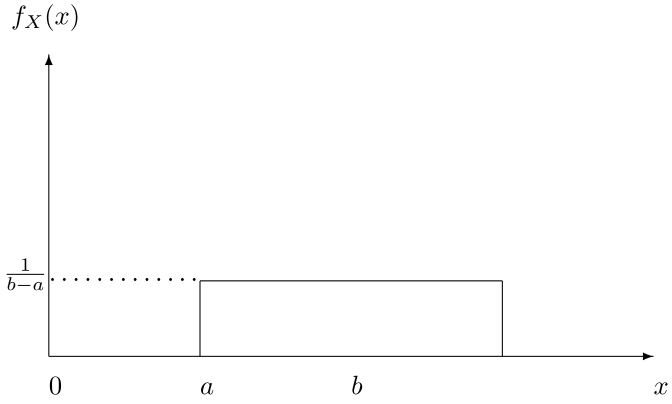


2 Some Important Random Variables

2.1 Uniform Random Variable on continuous interval $[a,b]$ $(a < b)$

- It is a Continuous Random Variable.
- $S_X = [a,b]$

- Its PDF is defined as:
  $f_X(x) = \begin{cases} \frac{1}{b-a} & \text{, if } x \in [a,b] \\ 0 & \text{, otherwise} \end{cases}$



- It is typically used to model random time/location in a bounded interval.

2.2 Bernoulli Random Variable

- It is a Discrete Random Variable.
- $S_X = \{0, 1\}$
- It may be defined using PMF as:

$$P_X(0) = p \text{ and } P_X(1) = 1 - p, \text{ for some } p \in [0, 1]$$

- It is typically used to model the outcome of a binary-outcome experience.

2.3 Binomial Random Variable

- It is a Discrete Random Variable.
- $S_X = \{0, 1, 2, ..., n\}$, for some n
- It may be specified via PMF as:

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \text{ for some p} \in [0, 1]$$

- It is typically used to model the number of "successes" in a sequence of n independent Bernoulli trials.

2.4 Geometric Random Variable

- It is a Discrete Random Variable.
- $S_X = \{0, 1, 2, ...\}$
- It may be specified via PMF as:

$$P_X(k) = (1-p)^{k-1}p, \text{ for some } p \in [0,1]$$

- It is typically used to model the number of independent Bernoulli trials until the first success is seen (including the last trial).

2.5 Exponential Random Variable

- It is a Continuous Random Variable.
- S_X is the set of all non-negative real numbers.
- It may be specified via PDF as:

$$f_X(k) = \lambda e^{-\lambda x},\ x \geq 0, \text{ for some given } \lambda > 0$$

- It is typically used to model waiting time, inter-arrival duration. (e.g., between customers, life time of a device, etc.)

2.6 Poisson Random Variable

- It is a Discrete Random Variable.
- $S_X = \{0, 1, 2, ...\}$
- It may be specified via PMF as:

$$P_X(k) = \frac{\alpha^k}{k!}\ e^{-\alpha}$$

- It is typically used to model the number of independent events happening during a given time interval or space.

A BIG EXAMPLE:

Suppose that the time now is 0 (hour), and I am waiting for phone calls.

Setting I: Suppose that John has said that he would call me before 2, and any time is equally likely for him to call.

Q1.1: *When will John call?*
Let x be the time at which John will call. Then x is a uniform Random Variable on the interval [0,2].

Setting II: Suppose phone calls arrive at random and independently. More precisely:

1. At every instant whether a phone call arrives is independent of whether a phone call would arrive at any other time instant.
2. A call would arrive at any time instant with equally likelyhood.

Now let's divide the time-axis into small intervals of length (Δ), and Δ is so small that within any interval there is at most one phone call that can arrive.

This translates to that the probability that a call hits every interval is the same. Suppose that this probability is P.

Q2.1: *Within a fixed interval (of length Δ), will a call arrive?*
The answer is a Bernoulli Random Variable. Remember that $S_X = \{0, 1\}$ for a Bernoulli Random Variable. In this example, one can assign 1 to the event of a call arrival and 0 for the event of no call arrival.

Q2.2: *In the time frame [0,2], how many intervals (of length Δ) will be hit by a phone call?*
The answer is a Binomial Random Variable with PMF:

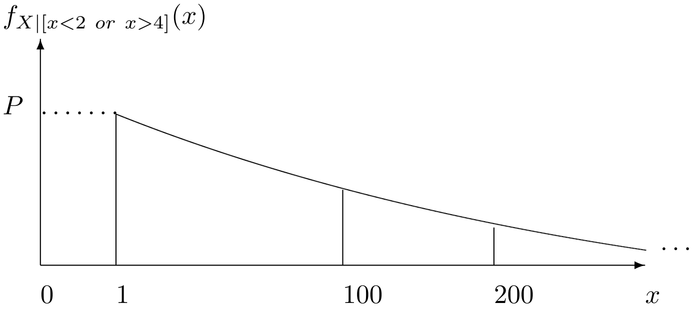
$$P_X(k) = \binom{2/\Delta}{k} p^k (1-p)^{(2/\Delta - k)}$$

Q2.3: *Which interval will be hit by a phone call first?*
The answer in terms of the interval number is a Geometric Random Variable with PMF:

$$P_X(k) = (1-p)^{k-1} p$$

Q2.3a: *Suppose that 100 time intervals (each of length Δ) have passed, and there were no phone calls in those intervals. Which interval is the first interval to be hit by a phone call?*
Let x be the interval number of the first interval hit by a phone call. We can write:

$$P[x > 100 + h \mid x > 100] = P[x > h], \text{ for any positive } h$$



Thus, we see a "memoryless" property associated with a Geometric Random Variable. In other words, the conditional characteristics after 100 time intervals have passed are exactly the same as those at time 0 (hour), but shifted after being scaled.

Q2.4: *Suppose in every time frame of length 2, there are α calls on average. Then, in the setting of Q2.2, we have:*

$$(\tfrac{2}{\Delta})p = \alpha$$

$$P_X(k) = \tbinom{n}{k}(\tfrac{\alpha}{n})^k(1-\tfrac{\alpha}{n})^{n-k}$$

where $[n = \frac{2}{\Delta}]$. Now, if we let $\Delta \to 0$, we get:

$$\lim_{\Delta\to 0} P_X(k) = \frac{\alpha^k}{k!}e^{-\alpha}$$

i.e. it is a Poisson distribution. We can conclude that Poisson distribution is a special case of exponential distribution.