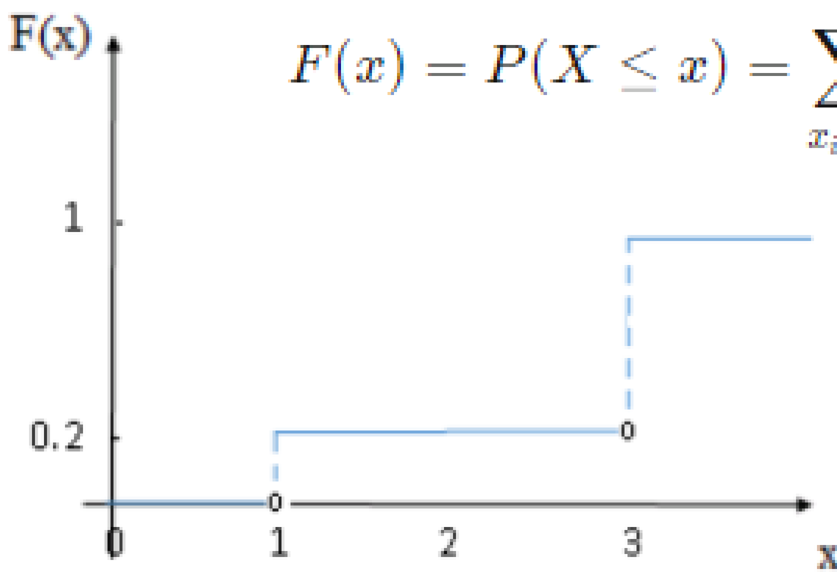
Random Variables

CDF of Random Variable

- Suppose that X is a random variable with values in $\mathbb{R}$. The Cumulative Distribution Function of X is the function $F:\mathbb{R}\rightarrow[0,1]$ is defined by $F(x)=P(X \le x)$, $x \in \mathbb{R}$

- For Discrete RV

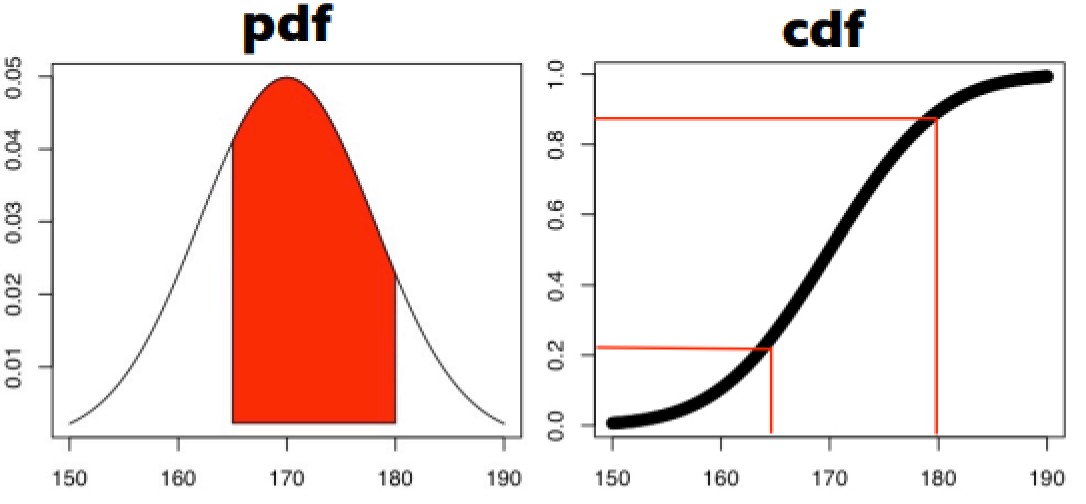
$$F(x) = P(X \le x) = \sum_{x_i \le x} P(X = x_i) = \sum_{x_i \le x} p(x_i),$$



CDF of Random Variable

- Suppose that X is a random variable with values in $\mathbb{R}$. The Cumulative Distribution Function of X is the function $F:\mathbb{R}\rightarrow[0,1]$ is defined by $F(x)=P(X \le x),\ x\in \mathbb{R}$

- **Continuous rv:** $F(x) = P(X \le x) = \int_{-\infty}^{x} f(y)dy$



$$P(a \le X \le b) = F(b) - F(a)$$

Expectation of Random Variable

- The expected value, or mean, of a random variable x denoted by E(x) or μ is a weighted average of the values the random variable may assume.
- In the discrete case the weights are given by the probability mass function, and in the continuous case the weights are given by the probability density function.
- The mean of a random variable may be interpreted as the average of the values assumed by the random variable in repeated trials of the experiment.

Expectation of Random Variable

- The expectation of a discrete random variable X is given by:

$$\mu = E(X) = \sum x f(x),$$ where *f(x)* is the PMF of x.

- The expectation of a continuous random variable X is given by:

$$\mu = E(X) = \int x\, f(x)\, dx,$$ where *f(x)* is the PDF of x.

Variance of Random Variable

- The variance of a random variable, denoted by Var(x) or σ^2, is a weighted average of the squared deviations from the mean.
  - For discrete RV: $\mathrm{Var}(x) = \sigma^2 = \Sigma(x - \mu)^2 f(x)$
  - For continuous RV: $\mathrm{Var}(x) = \sigma^2 = \int (x - \mu)^2 f(x)dx$

- The standard deviation, denoted σ, is the positive square root of the variance.

Questions

Q1. A service organization in a large town organizes a raffle each month. One thousand raffle tickets are sold for \$1 each. Each has an equal chance of winning. First prize is \$300, second prize is \$200, and third prize is \$100. Let, X denote the net gain from the purchase of one ticket.

- Construct the probability distribution of X
- Find the probability of winning any money in the purchase of one ticket.
- Find the expected value of X, and interpret its meaning.

Questions

a) If a ticket is selected as the first prize winner, the net gain to the purchaser is the \$300 prize less the \$1 that was paid for the ticket, hence $X=300-11=299$. There is one such ticket, so $P(299)=0.001$

Applying the same "income minus outgo" principle to the second and third prize winners and to the 997 losing tickets yields the probability distribution:

x	299	199	99	−1
P(x)	0.001	0.001	0.001	0.997

b) Let W denote the event that a ticket is selected to win one of the prizes. Using the table

$$P(W) = P(299)+P(199)+P(99) = 0.003$$

c) $E(X) = (299)\cdot(0.001) + (199)\cdot(0.001) + (99)\cdot(0.001) + (-1)\cdot(0.997)$

$= -0.4$

Questions

Q2. A discrete rv X has following probability distribution:

x	-1	0	1	4
P(x)	0.2	0.5	c	0.1

Find

a) c
b) P(0)
c) $P(X>0)$
d) $P(X\geq 0)$
e) $P(X\leq -2)$
f) The mean μ of X
g) The variance $\sigma 2$ of X
h) The standard deviation σ of X

Questions

Q2. A discrete rv X has following probability distribution:

x	-1	0	1	4
P(x)	0.2	0.5	c	0.1

Find

a) c	0.2
b) P(0)	0.5
c) $P(X>0)$	0.3
d) $P(X\geq 0)$	0.8
e) $P(X\leq -2)$	0
f) The mean μ of X	0.4
g) The variance σ^2 of X	1.84
h) The standard deviation σ of X	1.3565

Chebyshev's Inequality

- Chebyshev's inequality, is a theorem that characterizes the dispersion of a RV away from its mean (expectation).
- The probability that a random variable *x* deviates from its expected value (mean) μ on either side by at least k, for all k>0, is given as follows.

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

Median of Random Variable

- The median of the discrete random variable X, is the value of x for which P(X ≤ x) is greater than or equal to 0.5 and P(X ≥ x) is greater than or equal to 0.5.

Q3. Given the following probability density function of a discrete random variable, calculate the median of the distribution:

$$f(x) = \begin{cases} 0.2 & x = 1, 4 \\ 0.3 & x = 2, 3 \end{cases}$$

Median of Random Variable

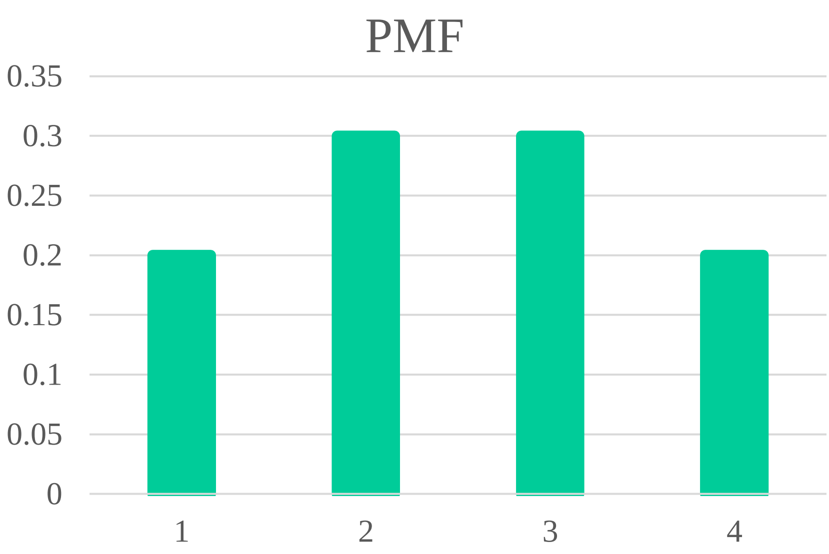
Q3. Given the following probability density function of a discrete random variable, calculate the median of the distribution:

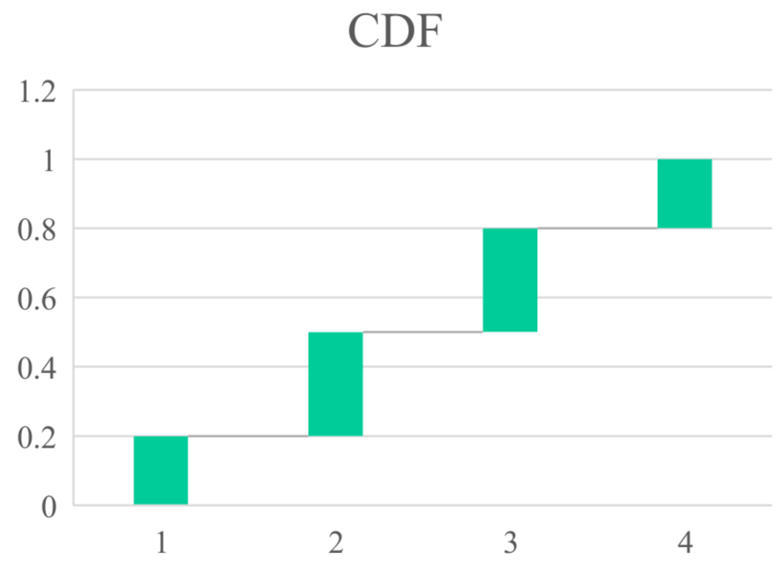
$$f(x) = \begin{cases} 0.2 & x = 1, 4 \\ 0.3 & x = 2, 3 \end{cases}$$

- **Solution:** The median of the distribution above is 2 because;

  $P(X \leq 2) = P(X=1) + P(X=2) = 0.2 + 0.3 = 0.5$ and,

  $P(X \geq 2) = P(X=2) + P(X=3) + P(X=4) = 0.3 + 0.3 + 0.2 = 0.8$





Median of Random Variable

- Let X be a continuous rv with probability density function, f(x). The median of X can be obtained by solving for c in the equation below:

$$\int_{-\infty}^{c} f(x)dx=0.5$$

- That is, it is the value for which the area under the curve from negative infinity to c is equal to 0.50.

Quantiles of Random Variable

- Let X is a real-valued random variable with Cumulative Distribution Function F.
- For $p \in (0,1)$, a value of x is called **a quantile of order *p*** for the distribution if

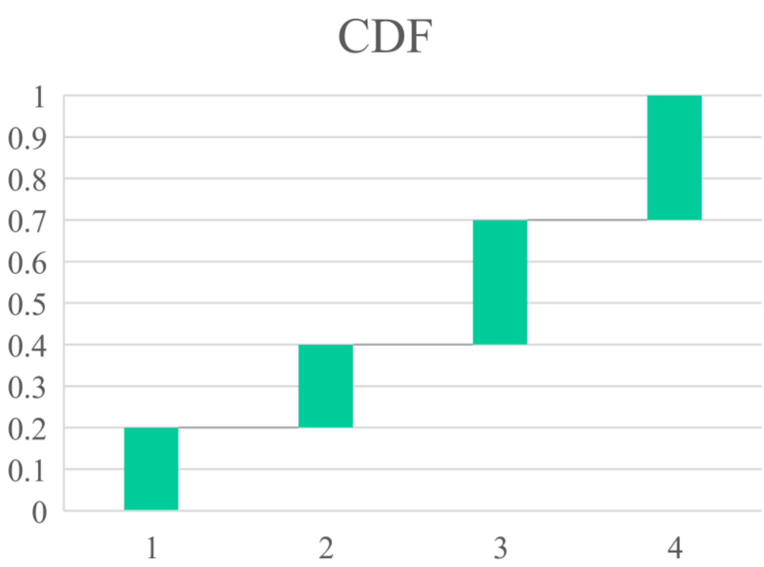
$$F(x-)=P(X < x) \leq p \text{ and } F(x)=P(X \leq x) \geq p \,.$$

- A quantile of order *p* is a value where the graph of the cumulative distribution function crosses *p*.
- Median is also called 50^{th} percentile.

Quantiles of Random Variable

- Given the following probability density function of a discrete random variable, calculate the 75th Percentile of the distribution:

$$f(x) = \begin{cases} 0.2 & x = 1, 4 \\ 0.3 & x = 2, 3 \end{cases}$$



Solution: The 75th percentile of the distribution is 4 because;

$$P(X < 3) = P(X=1)+P(X=2)$$
$$= 0.2+0.3 = 0.5 \quad \text{and,}$$
$$P(X \leq 3) = P(X=1)+P(X=2)+P(X=3)$$
$$= 0.2+0.3+0.3=0.8$$