DISCRETE PROBABILITY DISTRIBUTIONS

CHAPTER-3

CS6015-LINEAR ALGEBRA AND RANDOM PROCESSES

Sometimes the sum $S = \sum x f(x)$ does not converge absolutely, and the mean of the distribution does not exist. If $S = -\infty$ or $S = +\infty$, then we can sometimes speak of the mean as taking these values also. Of course, there exist distributions which do not have a mean value.

(12) Example. A distribution without a mean. Let X have mass function

$$f(k) = Ak^{-2} \quad \text{for} \quad k = \pm 1, \pm 2, \dots$$

where A is chosen so that $\sum f(k) = 1$. The sum $\sum_k k f(k) = A \sum_{k \neq 0} k^{-1}$ does not converge absolutely, because both the positive and the negative parts diverge. ●

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \in A^c, \end{cases}$$

and $\mathbb{E} I_A = \mathbb{P}(A)$.

(1) Example. Proofs of Lemma (1.3.4c, d). Note that

$$I_A + I_{A^c} = I_{A \cup A^c} = I_\Omega = 1$$

and that $I_{A\cap B} = I_A I_B$. Thus

$$\begin{aligned} I_{A\cup B} &= 1 - I_{(A\cup B)^c} = 1 - I_{A^c \cap B^c} \\ &= 1 - I_{A^c} I_{B^c} = 1 - (1 - I_A)(1 - I_B) \\ &= I_A + I_B - I_A I_B. \end{aligned}$$

Take expectations to obtain

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

More generally, if $B = \bigcup_{i=1}^n A_i$ then

$$I_B = 1 - \prod_{i=1}^n (1 - I_{A_i});$$

multiply this out and take expectations to obtain

(2) $$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_i \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \cdots + (-1)^{n+1} \mathbb{P}(A_1 \cap \cdots \cap A_n).$$

This very useful identity is known as the *inclusion–exclusion formula*.

Discrete Probability Distributions

- Bernoulli distribution
- Binomial distribution
- Trinomial distribution
- Poisson distribution
- Geometric distribution
- Negative binomial distribution

Bernoulli distribution

- A random variable *X* takes values 1 and 0 with probabilities *p* and *q* (= 1 - *p),* respectively.
- Sometimes we think of these values as representing the 'success' or the 'failure' of a trial.
- The mass function is

$$f(0) = 1 - p,\ f(1) = p,$$

- and it follows that $\boldsymbol{EX = p}$ and $\boldsymbol{var(X) = p(1 - p)}$.

Binomial Distribution

- We perform *n* independent Bernoulli trials $X_1, X_2, \ldots, X_n$ and count the total number of successes $Y = X_1 + X_2 + \ldots + X_n$.

- The mass function of *Y* is

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0,1,\ldots,n$$

- $\boldsymbol{EY = np}$ and $\boldsymbol{var(Y) = np(1-p)}$

Example

- A coin is tossed 10 times. What is the probability of getting exactly 6 heads?

Solution :

$$n = 10, p = 0.5, 1 - p = 0.5, x = 6$$

Using the formula from previous slide and substituting the above values we get

$$P(x = 6) =$$

Trinomial Distribution

- Suppose we conduct n trials, each of which results in one of three outcomes (red, white, or blue, say), where red occurs with probability p, white with probability q, and blue with probability $1 - p - q$. The probability of r reds, w whites, and $n - r - w$ blues is

$$\frac{n!}{r!w!(n-r-w)!}p^r q^w (1 - p - q)^{n-r-w}$$

this is the *trinomial distribution,* with parameters n, p, and q.

- The ‘multinomial distribution‘ is the obvious generalization of this distribution to the case of some number, say *t,* of possible outcomes.

Poisson Distribution

- A *Poisson* variable is a random variable with the Poisson mass function

$$f(k) = \frac{\lambda^k}{k!} e^{-\lambda}, \qquad k = 0,1,2, \dots$$

For some $\lambda > 0$

- Both the mean and the variance of this distribution are equal to λ.

Practice problems

1a. If calls to your cell phone are a Poisson process with a constant rate λ = 2 calls per hour, what's the probability that, if you forget to turn your phone off in a 1.5 hour movie, your phone rings during that time?

1b. How many phone calls do you expect to get during the movie?

https://view.officeapps.live.com/op/view.aspx?src=https://web.stanford.edu/~kcobb/hrp259/lecture5.ppt

Answer

1a. If calls to your cell phone are a Poisson process with a constant rate λ=2 calls per hour, what's the probability that, if you forget to turn your phone off in a 1.5 hour movie, your phone rings during that time?

$$\boldsymbol{f(k) = \frac{\lambda^k}{k!} e^{-\lambda}}$$

$X \sim$ Poisson (λ=2 calls/hour)

$P(X \geq 1) = 1 – P(X = 0)$

$P(X = 0) =$ [illegible]

$\therefore P(X \geq 1) = 1 – .05 \; = \; 95\%$ chance

1b. How many phone calls do you [illegible] the movie?

$E(X) \; =$ [illegible]

Geometric Distribution

- A *geometric* variable is a random variable with the geometric mass function

$$f(k) = p(1-p)^{k-1}, k = 1,2, \ldots$$

For some p in (0,1).

- Mean $= \frac{1}{p}$
- Variance $= \frac{1-p}{p^2}$

Negative Binomial Distribution

- $P(W_r = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, k = r, r+1, \ldots$
- The random variable W_r is the sum of *r* independent geometric variables. To see this, let X_1 be the waiting time for the first success, X_2 the *further* waiting time for the second success, X_3 the *further* waiting time for the third success, and so on. Then $\boldsymbol{X_1}$, $\boldsymbol{X_2}$, … are independent and geometric, and

$$W_r = X_1 + X_2 + \cdots + X_r$$

1. De Moivre trials. Each trial may result in any of t given outcomes, the ith outcome having probability p_i. Let N_i be the number of occurrences of the ith outcome in n independent trials. Show that

$$\mathbb{P}(N_i = n_i \text{ for } 1 \le i \le t) = \frac{n!}{n_1!\,n_2! \cdots n_t!}\, p_1^{n_1} p_2^{n_2} \cdots p_t^{n_t}$$

for any collection $n_1, n_2, \ldots, n_t$ of non-negative integers with sum n. The vector N is said to have the *multinomial distribution*.

(2) Definition. The **joint distribution function** $F : \mathbb{R}^2 \to [0, 1]$ of X and Y, where X and Y are discrete variables, is given by

$$F(x, y) = \mathbb{P}(X \leq x \text{ and } Y \leq y).$$

Their **joint mass function** $f : \mathbb{R}^2 \to [0, 1]$ is given by

$$f(x, y) = \mathbb{P}(X = x \text{ and } Y = y).$$

(3) Lemma. *The discrete random variables X and Y are independent if and only if*

(4) $$f_{X,Y}(x, y) = f_X(x) f_Y(y) \quad \textit{for all } x, y \in \mathbb{R}.$$

More generally, X and Y are independent if and only if $f_{X,Y}(x, y)$ can be factorized as the product $g(x)h(y)$ of a function of x alone and a function of y alone.

Read about Random Walks