STAT 511

Lecture 5: Discrete Random Variables, Distributions and Moments

Devore: Section 3.1-3.3

Prof. Michael Levine

January 23, 2019

Random Variables

- For a given sample space S, a random variable (RV) is any mapping $Y : S \to R$.
- Essentially, it is a function whose domain is the sample space and whose range is R.
- It is also possible to consider complex-valued random variables. We will not do it in the current course, however.

- A set is **denumerable** if and only if its elements can be placed in one-to-one correspondence with natural numbers.
- A set is **countable** if and only if it is either finite or denumerable.
- Example of a denumerable set : a set of all even natural numbers...Why? $2^1, 4^2, 6^3, 8^4, \ldots$
- Another example: a set of all integers...Indeed, $\ldots, -3^7, -2^5, -1^3, 0^1, 1^2, 2^4, 3^6, 4^8, \ldots$

- A discrete random variable is a RV whose possible values make up a countable sequence.
- Ex.(Discrete) A person attempts to log on to a power-sharing system; the outcome is either a success, coded by 1, or failure coded by zero. Thus, with $S = \{S, F\}$, we have

$$X(S) = 1, \ X(F) = 0.$$

- Any RV that only takes values 0 or 1 is called Bernoulli RV, in honor of Jacob Bernoulli $(1654 - 1705)$.

Example

- The quality control process: we sample batteries (or any other industrially manufactured product) as it comes off the conveyor line. Let F denote the faulty and S the good one. The sample space is $\mathcal{S} = \{S, FS, FFS, \ldots\}$. Let X be the number of batteries that is examined before the experiment stops. The, $X(S) = 1, X(FS) = 2, \ldots$.

Example of a non-discrete random variable

- Consider measuring the elevation above the sea level of a randomly chosen point within the continental US map (in feet).
- It will be

$$-282 \leq y \leq 14,494$$

where the left bound corresponds to the Death Valley and the right one to Mt. Whitney.
- This random variable is continuous.

Probability Distributions

- The probability distribution or probability mass function (pmf) of a discrete RV is defined for every number x as

$$p(x) = P(X = x) = P(\text{all } s \in S : X(s) = x)$$

- pmf specifies the probability of observing the value x when the experiment is performed
- It must satisfy:
  1. $p(x) \geq 0$
  2. $\sum_{\text{all possible } x} p(x) = 1$

Example

- Record preferences of a customer as

$$X = \begin{cases} 1 \text{ if laptop} \\ 0 \text{ if desktop} \end{cases}$$

- Assume that 20% of all customers selected a laptop. Then,

$$\begin{aligned} p(0) &= P(X = 0) = 0.8 \\ p(1) &= P(X = 1) = 0.2 \\ p(x) &= P(X = x) = 0 \text{ if } x \neq 0, 1 \end{aligned}$$

Cumulative distribution

- The cumulative distribution function (cdf) F(x) of a discrete RV variable X with pmf p(x) is

$$F(x) = P(X \leq x) = \sum_{y:y \leq x} p(y).$$

- It is the probability that X will be at most equal to x

Example I

- A store carries flash drives with either 1 GB, 2 GB, 4 GB, 8 GB, or 16 GB of memory
- The accompanying table gives the distribution of Y - the amount of memory in a purchased drive
-

y	1	2	4	8	16
p(y)	.05	.10	.35	.40	.10

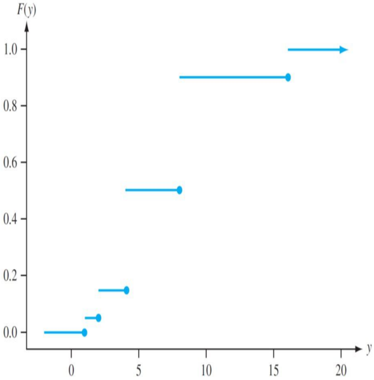
Example I

- E.g. $F(1) = P(Y \leq 1) = 0.5$ or $F(2) = P(Y \leq 2) = p(1) + p(2) = .15$
- But also $F(2.7) = P(Y \leq 2.7) = P(Y \leq 2) = .15$

$$F(y) = \begin{cases} 0 & y < 1 \\ .05 & 1 \leq y < 2 \\ .15 & 2 \leq y < 4 \\ .50 & 4 \leq y < 8 \\ .90 & 8 \leq y < 16 \\ 1 & 16 \leq y \end{cases}$$

Example I

- The representation of a CDF below is called a **step function**



Example II

- Starting at a fixed time, we observe the gender of each newborn child at a hospital until a boy is born. Let $p = P(B)$ and X the number of births observed until "success"
- Then,

$$p(x) = (1-p)^{x-1}p$$

for $x = 1, 2, 3 \dots$

- Verify that

$$F(x) = 1 - (1-p)^x$$

for any positive integer x

- More generally,

$$F(x) = \begin{cases} 0 & x \le 1 \\ 1 - (1-p)^{[x]} & x \ge 1 \end{cases}$$

where $[x]$ is the **integer part** of x

- As an example, if $p = 0.51$ $F(5) = 1 - (0.49)^5 = 0.9718$

Family of distributions

- Suppose that $p(x)$ depends on a parameter.
- Each value of the parameter determines a different probability distribution.
- Example I : a RV X is defined as $X = 1$ with prob. α and $X = 0$ with probability $1 - \alpha$. This is a whole *family* of distributions $p(x; \alpha)$. It is called a family of Bernoulli distributions
- Example II: a RV Y that is defined as the number of "failures " before the first success with each trial being independent and the probability of "success" being p; see the example above.
- Such a family is called the family of **geometric** RV's with the parameter p

Proposition

- For any two numbers a and b with

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

  where $a-$ represents the largest possible X value that is strictly less than a.

- If a and b are integers,

$$P(a \leq X \leq b) = F(b) - F(a-1)$$

Example

- Let X be the number of days of sick leave taken by a randomly selected employee of a large company during a particular year
- If the maximum number of allowable sick days per year is 14, possible values of X are 0, 1, . . . , 14
- Check that $F(0) = .58$, $F(1) = .72$, $F(2) = .76$, $F(3) = .81$, $F(4) = .88$, and $F(5) = .94$
- E.g.

$$P(2 \leq X \leq 5) = P(X = 2, 3, 4, \text{ or } 5) = F(5) - F(1) = .22$$

- Or,

$$P(X = 3) = F(3) - F(2) = 0.05$$

Chevalier de Méré - Pascal-Fermat problem

- How to split the pot of an interrupted dice game? Let each of the two players select a number from the set $S = \{1, 2, 3, 4, 5, 6\}$
- For each roll of a fair die that produces one of their respective numbers, the corresponding player receives a token; the one who accumulates 5 tokens, receives 100
- What if the game is interrupted when Player A has 4 tokens and the Player B just one?
- The probability that Player B would have won the pot is that his number appears 4 more times before A's number appears one more time...This is $0.5^4 = 0.0625$
- According to Pascal and Fermat, B is entitled to $0.0625 * 100 = 6.25$ from the pot and the remaining 93.75 go to Player A

Expected value

- For a random variable X with pmf $p(x)$, the expected (mean) value is
$$E(X) = \sum_{x \in D} xp(x)$$
where D is the set of all possible values x.
- Consider a slot machine that pays a jackpot of 1000 with $p = 0.0005$ and otherwise nothing.
- Define $X = 1000$ with $p = 0.0005$ and $X = 0$ with $1 - p = 0.9995$.
- The **fair value** is $E\,X = 1000 * 0.0005 + 0 * 0.9995 = 0.5-$ this is what you should charge for participating in this game

Examples of lotteries

- The win is 10 if the fair coin produces H or 0 otherwise; if you don't pay, you receive 5.
- The fair value is $10 * 0.5 = 5-$ a "rational" person should be indifferent...are you?
- What if the win is 10000 and you have to pay 5000 to participate? Or 2 million and 1 million?
- Now imagine that you either receive 1 million for sure or get 5 million with $p = 0.5$ (tossing the fair coin). The fair value is 2.5 million...but do you REALLY want to play the game?

Credit card example

- Let X be the number of credit cards a student carries.

x	$p(x)$
0	0.08
1	0.28
2	0.38
3	0.16
4	0.06
5	0.03
6	0.01

$$E(X) = x_1 p_1 + \ldots + x_n p_n = 0*(0.08) + 1*(0.28) + \ldots + 6*(0.01) = 1.97$$

Expected value of a Bernoulli random variable

- **Example** Consider a Bernoulli RV with pmf

$$p(x) = \begin{cases} 1-p & \text{if} \quad \text{x=0} \\ p & \text{if} \quad \text{x=1} \\ 0 & \text{if} \quad \text{otherwise} \end{cases}$$

$$E(X) = 0 * p(0) + 1 * p(1) = p$$

- The expected value of X is the probability of success.
- Note that the expected value is the weighted mean and not just the simple mean of outcomes which is equal to 0.5 regardless of the values of p

Expected value of a geometric random variable

- **Example** Let X now be the number of children born up to and including the first boy is

$$p(x) = p(1-p)^{x-1}$$

for $x = 1, 2, 3, \ldots$.

$$E(X) = \sum_{x \in D} x \cdot p(x) = \sum_{x=1}^{\infty} xp(1-p)^{x-1} = \frac{1}{p}$$

Expected value of a function and a St. Petersburg paradox

- If the RV X has the set of possible values D and pmf $p(x)$, then the expected value of any function $h(X)$ is denoted

$$E(h(X)) = \sum_{x \in D} h(x) \cdot p(x)$$

- Let the jackpot start at 1 and double each time T is observed at the toss of a fair coin. When H is observed, the game is terminated
- How much would you pay for the privilege of playing this game? How much would you charge if you were responsible for making the payoff?

St. Petersburg paradox

- There is a very small possibility of a large payoff; most people answer that they won't pay more than 4 for the privilege of paying this game
- Most people do request a fairly large payment, recognizing the possibility of the large payoff
- If X is the number of tails observed until the end of the game,

$$f(x) = P(x \text{ consecutive T's}) = 0.5^x$$

- The payoff is $Y = 2^X$ and

$$E\,Y = \sum_{x=0}^{+\infty} 2^x * 0.5^x = \infty$$

- In this case, the "fair value" provides very little insight into how much you'd want to pay to take part in this game

Expected Value Properties

- For any two real numbers a and b

$$E(aX + b) = a \cdot E(X) + b$$

- Corollaries:
  1. $$E(aX) = a \cdot E(X)$$
  2. $$E(X + b) = E(X) + b$$
- Let X be a discrete random variable such that $P(X = c) = 1$. Then, $E\,X = c$.

Variance

- Let X have pmf p(x), and expected value μ. Then the variance of X is

$$V(X) = \sum_{x \in D} (x - \mu)^2 \cdot p(x) = E[(X - \mu)^2]$$

- Alternative notations: σ_X^2 or σ^2
- The standard deviation of X is

$$\sigma_X = \sqrt{\sigma_X^2}$$

- σ_X^2 is the **population variance**
- An alternative formula is $Var\, X = E\, X^2 - (E\, X)^2$

Examples

1. Variance of a Bernoulli random variable is $Var\ X = E(X-\mu)^2 = (-\mu)^2 * (1-p) + (1-\mu)^2 * p = p^2(1-p) + (1-p)^2 p = p(1-p)$
2. The quiz scores for a particular student are given as 22, 25, 20, 18, 12, 24, 20, 20, 25, 24, 25, 18.

Value	12	18	20	22	24	25
Frequency	1	2	4	1	2	3
Probability	.08	.15	.31	.08	.15	.23

2a. $\mu = 0.08 * 12 + \ldots + 0.23 * 25 = 21$

2b.

$$V(X) = p_1(x_1-\mu)^2 + p_2(x_2-\mu)^2 + \ldots + p_n(x_n-\mu)^2$$

$$V(X) = .08\cdot(12-21)^2 + .15\cdot(18-21)^2 + \ldots + .23\cdot(25-21)^2 = 13.25$$