Stat 201: Introduction to Statistics

Standard 18: Probability Distributions – Discrete Distributions

Random Variable

- **Random Variable –** a numerical measurement of the outcome of a random phenomena
  - Capital letters refer to the random variable
  - Lower case letters refer to specific realizations

- Recall our definitions of Discrete and Continuous quantitative variables from before

Random Variable

- **Discrete Example:** Number of goals in an EPL soccer match
  - We refer to the number of goals in an EPL soccer match as X, until we have a concrete observation
  - x=2 goals is a realization – a concrete observation

Random Variable

- **Continuous Example:** Height of Americans
  - We refer to the Height of Americans as X, until we have a concrete observation
  - x=72 inches is a realization – a concrete observation

Discrete Distributions!

- **Probability Distribution –** a summary of all possible outcomes of a random phenomena along with their probabilities
  - **Example 1**: Number of goals scored in an EPL game
  - **Example 2&3:** Number of red lights on your way to work
  - **Example 4:** Number of free throws made

Random Variable: Discrete

- The possible outcomes must be countable
  - Remember quantitative discrete variables from chapter 2? If not, you should look back!
- We have a **valid** discrete probability distribution if
  1. Our outcomes are discrete (countable)
  2. All the probabilities are valid
     - $0 \leq P(x) \leq 1 \; for \; all \; outcomes \; x$
  3. We've accounted for all possible outcomes
     - $\sum P(x) = 1$

Example 1: Discrete Distributions

- **Example**: number of goals scored in an EPL soccer match
- The number of goals is countable
- All probabilities are between 0 and 1
- $\sum P(x) = .0711 + .1974 + .2158 + .1842 + .1658 + .1026 + .0447 + .0105 + .0026 + .0053 = 1$

X = # of Goals	P(x) = Probability
0	.0711
1	.1974
2	.2158
3	.1842
4	.1658
5	.1026
6	.0447
7	.0105
8	.0026
9	.0053
TOTAL	1

Example 2: Discrete Distribution

- Example: Number of red lights on the way to work (there are only three red lights on your way to work – this means you can catch 0,1,2 or 3 lights on your way to work.)

X = Number of lights	P(x) = Probability
0	.10
1	.10
2	.10
3	.40

- The number of red lights is countable
- All probabilities are between 0 and 1
- $\sum P(x) = .10 + .10 + .10 + .40 = .70$
- Since $\sum P(x) = .70 \neq 1$ we **do not** have a valid Discrete Dist.

Example 3 Discrete Distributions Route 2

- Example: Number of red lights on the way to work (there are only three red lights on your way to work – this means you can catch 0,1,2 or 3 lights on your way to work.)

X = Number of lights	P(x) = Probability
0	.40
1	.30
2	.20
3	.10

- The number of red lights is countable
- All probabilities are between 0 and 1
- $\sum P(x) = .40 + .30 + .20 + .10 = 1$

Example 3 Discrete Distributions Route 2

- Example: Number of red lights on the way to work (there are only three red lights on your way to work – this means you can catch 0,1,2 or 3 lights on your way to work.)

X = Number of lights	P(x) = Probability
0	.20
1	.30
2	.10
3	.40

- The number of red lights is countable
- All probabilities are between 0 and 1
- $\sum P(x) = .20 + .30 + .10 + .40 = 1$

Example 4: Discrete Distribution

- **Example:** Number of free throws made by a basketball player in 2 tries

X = Number Made	P(x) = Probability
0	.40
1	.40
2	.20

- The number of free throws is countable
- All probabilities are between 0 and 1
- $\sum P(x) = .40 + .40 + .20 = 1$

The Mean of a Discrete Distribution

- The **mean** of a probability distribution represents the average of a large number of observed values. [Remember: in the long run]

- We denote this with the Greek letter as below
  - $\mu_x = E(X) = Expected\ value\ of\ x = \sum xP(x)$

Example 1: Discrete Distributions

- **Example**: # of goals scored in an EPL soccer match
- $\mu_x = E(x) = \sum x * P(x) = 0 + .1974 + .4316 + .5526 + .6632 + .5130 + .2682 + .0735 + .0208 + .0477 = 2.768$

X = # of Goals	P(X)	X*P(X)
0	.0711	0*.0711=0
1	.1974	1*.1974=.1974
2	.2158	2*.2158=.4316
3	.1842	3*.1842=.5526
4	.1658	4*.1658=.6632
5	.1026	5*.1026=.5130
6	.0447	6*.0447=.2682
7	.0105	7*.0105=.0735
8	.0026	8*.0026=.0208
9	.0053	9*.0053=.0477
TOTAL	1	2.768

Example 1: Discrete Distributions

- **Example**: # of goals scored in an EPL soccer match
- $\mu_x = E(x) = \sum x * P(x) = 0 + .1974 + .4316 + .5526 + .6632 + .5130 + .2682 + .0735 + .0208 + .0477 = 2.768$
- We like to write the interpretation in reasonable terms
  - **"On average, we expect** between two and three goals in an EPL soccer match"

Example 2 Discrete Distributions Comparing Routes: Route 1

X = Number of lights	P(X)	X*P(X)
0	.40	0*.40=0
1	.30	1*.30=.30
2	.20	2*.20=.40
3	.10	3*.10=.30

- $E(X) = \sum xP(x) = 0 + .3 + .4 + .3 = 1$
- **"On average, we expect** that Route 1 will result in hitting one red light"

Example 3 Discrete Distributions Comparing Routes: Route 2

X = Number of lights	P(X)	X*P(X)
0	.20	0*.20=0
1	.30	1*.30=.30
2	.10	2*.10=.20
3	.40	3*.40=1.20

- $E(X) = \sum xP(x) = 0 + .30 + .20 + 1.20 = 1.7$
- **“On average, we expect** that Route 2 will result in hitting between one and two red lights”

Example 2&3 Discrete Distributions Comparing Routes

- Route 1
  - $E(X) = \sum xP(x) = 1$
- Route 2
  - $E(X) = \sum xP(x) = 1.7$

- Route 2 will result in more lights **on average**

Example 4: Discrete Distribution

- **Example:** Number of free throws made by a basketball player in 2 tries

X = Number Made	P(x) = Probability	x*P(x)
0	.40	0*.40 = 0
1	.40	1*.40 = .40
2	.20	2*.20 = .40

- $\mu_x = E(x)$

$$= \sum x * P(x) = 0 + .40 + .40 = .80$$

- **“On average, we expect** between zero and one free throw in two tries”

Why don’t I get this?

- Probabilities and expected values are much different than what we did in Chapter 2 where you found the sample mean by adding up values and dividing.
- Expected value in the sense of the discrete distribution is what we would expect to see on average if we completed the random experiment infinitely many times
  - i.e. if I took the same route to work every day for the rest of time how many lights would I expect to see on average over all of those trips.

Discrete Distributions on YouTube

- Introduction
  - https://www.youtube.com/watch?v=mrCxwEZ_22o

- Mean and standard deviation of a Discrete Distribution on your TI calculator
  - https://www.youtube.com/watch?v=7l9_s5Zj9gQ

Discrete Distributions on your TI

- INPUT:
  - Press STAT
  - Press ENTER with ‘Edit’ highlighted
  - Enter the X data into the L1 column
  - Enter the P(X) or frequency data into the L2 column
  - Press STAT
  - Press → to CALC
  - Press ENTER with ‘1: 1-Var Stats’ highlighted
  - Press 2^{nd}
  - Press 1
  - Press ,
  - Press 2^{nd}
  - Press ,
  - Press Enter

Discrete Distributions on your TI

- OUTPUT:
  - We don’t care about all of the output here – the important ones are listed below
  - $\bar{x}$ = the mean of the discrete probability distribution
  - σ_x = the standard deviation of this discrete probability distribution

Random Variables

Random Variable	A numerical measurement of the outcome of a random phenomena -Capital letters refer to the random variable -Lower case letters refer to specific realizations
Categorical Random Variable	Random variables that belong to a set of categories
Quantitative Random Variable	Random variables that take on numerical values

Discrete Distribution

Probability Distribution	A summary of all possible outcomes of a random phenomena along with their probabilities
Valid discrete probability distribution	IF: 1. Our outcomes are discrete (countable) 2. All the probabilities are valid $0 \leq P(x) \leq 1 \; for \; all \; outcomes \; x$ 3. We've accounted for all possible outcomes $\sum P(x) = 1$
Expected value of discrete X	$\mu_x = E(X) = Expected \; value \; of \; x$ $= \sum xP(x)$