Stats HW2 PS02

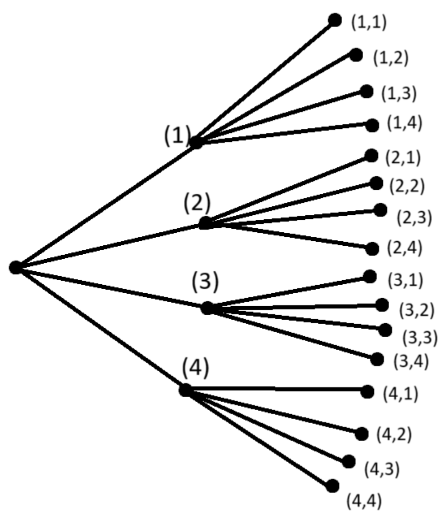
Sejal Nimkar

2024-09-17

1.Use the urn {1,1,1,1,2,2,2,3,3,4}. The experiment is to draw two tickets with replacement and let the random variable X be the sum of these two tickets.

a. Find the PMF of X.

Ans: Let's draw a decision tree for the following experiment.



We will derive $X(S)$ using the above decision tree.

Total number of outcomes = 16

X is the random variable that assigns the sum of two tickets to every outcome. Thus $X(S)$ is:

$X(S) = \{2, 3, 4, 5, 6, 7, 8\}$.

Since we are drawing tickets with replacement, every outcome has an equally likely chance of occurring.

Therefore PMF of all values of X can be given as $f(x) = P(X = x)$:

$f(2) = P(X = 2) = P(\{(1, 1)\}) = \dfrac{1}{16} = 0.0625$

$f(3) = P(X = 3) = P(\{(1, 2), (2, 1)\}) = \dfrac{2}{16} = 0.125$

$f(4) = P(X = 4) = P(\{(1, 3), (2, 2), (3, 1)\}) = \dfrac{3}{16} = 0.1875$

$f(5) = P(X = 5) = P(\{(1,4),(2,3),(3,2),(4,1)\}) = \frac{4}{16} = 0.25$

$f(6) = P(X = 6) = P(\{(2,4),(3,3),(4,2)\}) = \frac{3}{16} = 0.1875$

$f(7) = P(X = 7) = P(\{(3,4),(4,3)\}) = \frac{2}{16} = 0.125$

$f(8) = P(X = 8) = P(\{(4,4)\}) = \frac{1}{16} = 0.0625$

b. Find the CDF of X.

Ans: The CDF i.e. the cumulative distributive function can be calculated by using the PMF of X calculated in the earlier question:

CDF is denoted as $F(x) = P(X \leq x)$. Therefore

$F(2) = P(X \leq 2) = \frac{1}{16} = 0.0625$

$F(3) = P(X \leq 3) = \frac{3}{16} = 0.1875$

$F(4) = P(X \leq 4) = \frac{6}{16} = 0.375$

$F(5) = P(X \leq 5) = \frac{10}{16} = 0.625$

$F(6) = P(X \leq 6) = \frac{13}{16} = 0.8125$

$F(7) = P(X \leq 7) = \frac{15}{16} = 0.9375$

$F(8) = P(X \leq 8) = \frac{16}{16} = 1$

c. Find the expected value and variance of X.

Ans: The equation for expected value is given as:

$E(X) = \sum_{x \in X(S)} xP(X = x) = \sum_{x \in X(S)} xf(x)$

Thus $E(X) = \sum_{x \in X(S)} xP(X = x) = (2)(0.0625) + (3)(0.125) + (4)(0.1875) + (5)(0.25) + (6)(0.1875) + (7)(0.125) + (8)(0.0625)$

$\therefore E(X) = \mu = 5$

The equation for variance of X is given by:

$\text{Var}(X) = E[(X - \mu)^2] = \sum_{x \in X(S)} (x - \mu)^2 f(x).$

Thus $\text{Var}(X) = (2-5)^2(0.0625) + (3-5)^2(0.125) + (4-5)^2(0.1875) + (5-5)^2(0.25) + (6-5)^2(0.1875) + (7-5)^2(0.125) + (8-5)^2(0.0625)$

$\therefore \text{Var}(X) = (9)(0.0625) + (4)(0.125) + (1)(0.1875) + (0)(0.25) + (1)(0.1875) + (4)(0.125) + (9)(0.0625)$

$\therefore \text{Var}(X) = 2.5$

d. What is F(3),P(X > 6) and P(5 < X < 8). Ans: The cumulative distribution function $F(y) = P(X \leq y)$ for random variable X for $x = 3$ is given as :

$F(3) = P(X \leq 3) = \frac{3}{16} = 0.1875$

To calculate $P(X > 6)$, we can use the following equation:

$P(X > 6) = 1 - P(X \leq 6)$

$\therefore P(X > 6) = 1 - 0.8125(fromQ1b)$

$\therefore P(X > 6) = 0.1875$

And to calculate $P(5 < X < 8)$, we can use either of the equations:

$P(5 < X < 8) = F(8) - F(5) - f(8)$ or$P(5 < X < 8) = f(6) + f(7)$

$\therefore P(5 < X < 8) = 0.1875 + 0.125 = 0.3125$

2. Toss three fair coins. Let X be the random variable that uses the rule of assignment: $(2 \times$ number of tails$)-$number of heads. For example if the outcome is TTH then $X(TTH) = 2 \times 2 - 1 = 3$. Determine each of the following:

a. The sample space S.

Ans: The sample space of S will include all the outcomes that can occur for the experiment with three coin tosses. Denoting H as the number of heads and T as the number of tails, the sample space for the experiment is as follows:

$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

b. The range of X.

Ans: The values that X assigns to each outcome is given by $(2 \times numberoftails) - numberofheads$. Thus for each outcome, the value can be calculated as follows:

a) For number of heads = 3, number of tails = 0 :

$X(HHH) = (2 \times 0) - 3 = 0 - 3 = -3$

b) For number of heads = 2, number of tails = 1 ;

$X(HHT) = X(HTH) = X(THH) = (2 \times 1) - 2 = 2 - 2 = 0$

c) For number of heads = 1, number of tails = 2;

$X(HTT) = X(THT) = X(TTH) = (2 \times 2) - 1 = 4 - 1 = 3$

d) For number of heads = 0, number of tails = 3 ;

$X(HHT) = X(HTH) = X(THH) = (2 \times 3) - 0 = 6 - 0 = 6$

Thus range of X is :

$X(S) = \{-3, 0, 3, 6\}$

c. The CDF of X.

Ans: $F(y) = P(X \leq y)$

Let's calculate probability mass function for all the values belonging to the range of X. Since the coin toss involves three fair coins, each outcome is equally likely i.e. every outcome has an equal probability of occurring. Thus the probability of every outcome can be calculated using:

$f(x) = P(X = x) = \dfrac{\text{number of successes (x)}}{\text{total number of outcomes}}$

Total number of outcomes = 8

$\therefore$ for value of X = -3, probability is:

$f(-3) = P(X = -3) = P(\{HHH\}) = \dfrac{1}{8}$

for value of X = 0, probability is:

$f(0) = P(X = 0) = P(\{HHT, HTH, THH\}) = \frac{3}{8}$

for value of X = 3, probability is:

$f(3) = P(X = 3) = P(\{HTT, THT, TTH\}) = \frac{3}{8}$

for value of X = 6, probability is:

$f(6) = P(X = 6) = P(\{TTT\}) = \frac{1}{8}$

Thus the cdf for X can be given as follows:

If $y < -3$ for e.g. $y = -4.5$ or $y = -19.6$:

then $F(y) = P(X \leq y) = P(X \leq -3) = \frac{1}{8} = 0.125$

If $y \in (-3, 0)$ for e.g $y = -2.5$ or $y = -0.3$:

then $F(y) = P(X \leq y) = P(X \leq 0) = \frac{4}{8} = 0.5$

If $y \in (0, 3)$ for e.g $y = 1.6$ or $y = 2.7$:

$F(y) = P(X \leq y) = P(X \leq 3) = \frac{7}{8} = 0.875$

If $y \in (3, 6)$ for e.g $y = 4.8$ or $y = 5$:

$F(y) = P(X \leq y) = P(X \leq 6) = \frac{8}{8} = 1$

d. The PMF of X.

Ans: As calculated in question 2b, we have the PMF for every value assigned by X to every outcome as follows:

$f(-3) = P(X = -3) = P(\{HHH\}) = \frac{1}{8} = 0.125$

$f(0) = P(X = 0) = P(\{HHT, HTH, THH\}) = \frac{3}{8} = 0.375$

$f(3) = P(X = 3) = P(\{HTT, THT, TTH\}) = \frac{3}{8} = 0.375$

$f(6) = P(X = 6) = P(\{TTT\}) = \frac{1}{8} = 0.125$

e. The expected value of X.

Ans: The equation for expected value is given as:

$$E(X) = \sum_{x \in X(S)} xP(X = x) = \sum_{x \in X(S)} xf(x)$$

Thus $E(X) = \sum_{x \in X(S)} xP(X = x) = (-3)(0.125) + (0)(0.375) + (3)(0.375) + (6)(0.125)$

$\therefore E(X) = \mu = 1.5$

f. The variance and standard deviation of X.

Ans: The equation for variance of X is given by:

$\mathrm{Var}(X) = E[(X - \mu)^2] = \sum_{x \in X(S)} (x - \mu)^2 f(x).$

As derived in the earlier question, value of μ i.e. the expected value is 1.5.

Thus variance can be given as:

$\mathrm{Var}(X) = (-3 - 1.5)^2(0.125) + (0 - 1.5)^2(0.375) + (3 - 1.5)^2(0.375) +$

$(6 - 1.5)^2(0.125)$

$\therefore \mathrm{Var}(X) = (20.25)(0.125) + (2.25)(0.375) + (2.25)(0.375) + (20.25)(0.125)$

$\therefore \mathrm{Var}(X) = 6.75$

Thus variance of X is 6.75.

Standard deviation is denoted by:

$\sigma = \sqrt{\mathrm{Var}(X)} = \sqrt{E[(X - \mu)^2]} = \sqrt{\sum_{x \in X(S)} (x - \mu)^2 f(x)}.$

$\therefore \sqrt{\mathrm{Var}(X)} = \sqrt{6.75} = 2.598$

Thus standard deviation of X is 2.598.

3. The Association for Research and Enlightenment (ARE) in Virginia Beach, VA, offers daily demonstrations of a standard technique for testing extrasensory perception (ESP). A "sender" is seated before a box on which one of five symbols (plus, square, star, circle, wave) can be illuminated. A random mechanism selects symbols in such a way that each symbol is equally likely to be illuminated. When a symbol is illuminated, the sender concentrates on it and a "receiver" attempts to identify which symbol has been selected. The receiver indicates a symbol on the receiver's box, which sends a signal to the sender's box that cues it to select and illuminate another symbol. This process of illuminating, sending, and receiving a symbol is repeated 25 times. Each selection of a symbol to be illuminated is independent of the others. The receiver's score (for a set of 25 trials) is the number of symbols that s/he correctly identifies. For the purpose of this exercise, please suppose that ESP does not exist.

a) How many symbols should we expect the receiver to identify correctly?

Ans: This experiment can be analysed as a binomial experiment where:

a. Total number of trials (n)= 25
b. There are five possible symbols in each trial.
c. Probability (p) of a symbol being correctly guessed is $\frac{1}{5} = 0.2$

For a binomial experiment, the expected number of successes is given by :

$E(X) = n \times p$

Thus the expected number of successes is :

$E(X) = 25 \times 0.2 = 5$

b. The ARE considers a score of more than 7 matches to be indicative of ESP. What is the probability that the receiver will provide such an indication?

Ans: The probability that the receiver will score more than 7 symbols can be given as :

$P(X > 7) = 1 - P(X \leq 7)$

To calculate $P(X \leq 7)$, we can use cumulative distributive function with R:

```
pbinom(q = 7,size= 25, prob = 0.2)
```

```
## [1] 0.8908772
```

Thus the value of $P(X > 7)$ is:

```
print(1 - pbinom(q=7,size=25,prob=0.2))
```

```
## [1] 0.1091228
```

Thus the probability of the receiver scoring the symbols more than 7 correctly is 0.109.

c. The ARE provides all audience members with scoring sheets and invites them to act as receivers . Suppose that, as on August 31, 2002, there are 21 people in attendance: 1 volunteer sender, 1 volunteer receiver, and 19 additional receivers in the audience. What is the probability that at least one of the 20 receivers will attain a score indicative of ESP?

Ans: The probability that a single receiver scores more than 7 symbols correctly is 0.109.

The probability that the receiver **does not** score more than 7 symbols correctly is $1 - P(X > 7)$

$\therefore 1 - P(X > 7) = P(X \leq 7) = 0.8908$ (as determined above).

The probability that out of 20 receivers, no receiver is able to score more than 7 symbols correctly is given as

$P(X \leq 7)^{20}$ i.e. $0.8909^2 0$

Thus, probability of at least one receiver being able to score more than 7, which will be indicative of ESP can be given as:

$1 - P(X \leq 7)^{20}$ i.e. $1 - 0.8908^2 0 =$

```
print(1 - 0.8908^20)
```

```
## [1] 0.9010071
```

Thus the probability of at least one receiver being able to score more than 7 is 90%.

4. Mike teaches two sections of Applied Statistics each year for thirty years, for a total of 1500 students. Each of his students spins a penny 89 times and counts the number of Heads . Assuming that each of these 1500 pennies has $P(Heads) = 0.3$ for a single spin, what is the probability that Mike will encounter at least one student who observes no more than two Heads ?

Ans: Every student will toss 89 coins where probability of getting a head is 0.3.

Thus the probability of a single student observing no more than two heads is: $P(X \leq 2)$, which equals to:

$P(X \leq 2)$ = (P({H =0})+P({H = 1})+P({H = 2})

$\therefore P(\{H = 0\}) = \binom{89}{0} 0.3^0 (1 - 0.3)^{89-0} = 1.635 \times 10^{-14}$

$\therefore P(\{H = 1\}) = \binom{89}{1} 0.3^1 (1 - 0.3)^{89-1} = 6.239E \times 10^{-13}$

$\therefore P(\{H = 2\}) = \binom{89}{2} 0.3^2 (1 - 0.3)^{89-2} = 1.176E \times 10^{-11}$

Since the values obtained are too small, we shall represent the sum of these values i.e $P(X \leq 2)$ as p.

The probability that a single student observes no more than two heads is p.

The probability that a single student observes more than two heads is $1 - p$

The probability that none of the 1500 students observes no more than two heads is :

$(1 - p)^{1500}$.

Therefore, the probability that at least one of the 1500 students observes no more than 2 Heads is:

$1 - (1 - p)^{1500}$