

# S520 Homework 3

Jack McShane

2022-02-19

## Problem 1

- c) **Determine the expected value of X.**

The expected value for a random variable can be calculated by taking a weighted sum where each possible value of the variable is weighted according to its probability. This can be done in R as so:

```
vals <- c(1, 3, 4, 6)
probs <- c(.1, .4, .4, .1)
expectation <- sum(vals*probs)
```

```
## [1] "Expected value: 3.5"
```

- d) **Determine the variance of X.**

The variance of a random variable measures how much observed values differ from the mean or expected value of the variable. It can be calculated using R in this manner:

```
squared_diff <- (vals - expectation)^2
variance <- sum(squared_diff * probs)
```

```
## [1] "Variance of X: 1.45"
```

- e) **Determine the standard deviation of X.**

The standard deviation of a random variable's distribution indicates how spread values of the random variable are from the mean. It can be calculated as so:

```
stddev <- sqrt(variance)
```

```
## [1] "Standard Deviation of X: 1.20415945787923"
```

## Problem 2

- a) **Determine the pmf of X.**

The pmf can be calculated as shown below. The `get_probs` function calculates the probability of each value of X according to the equation given by the problem.

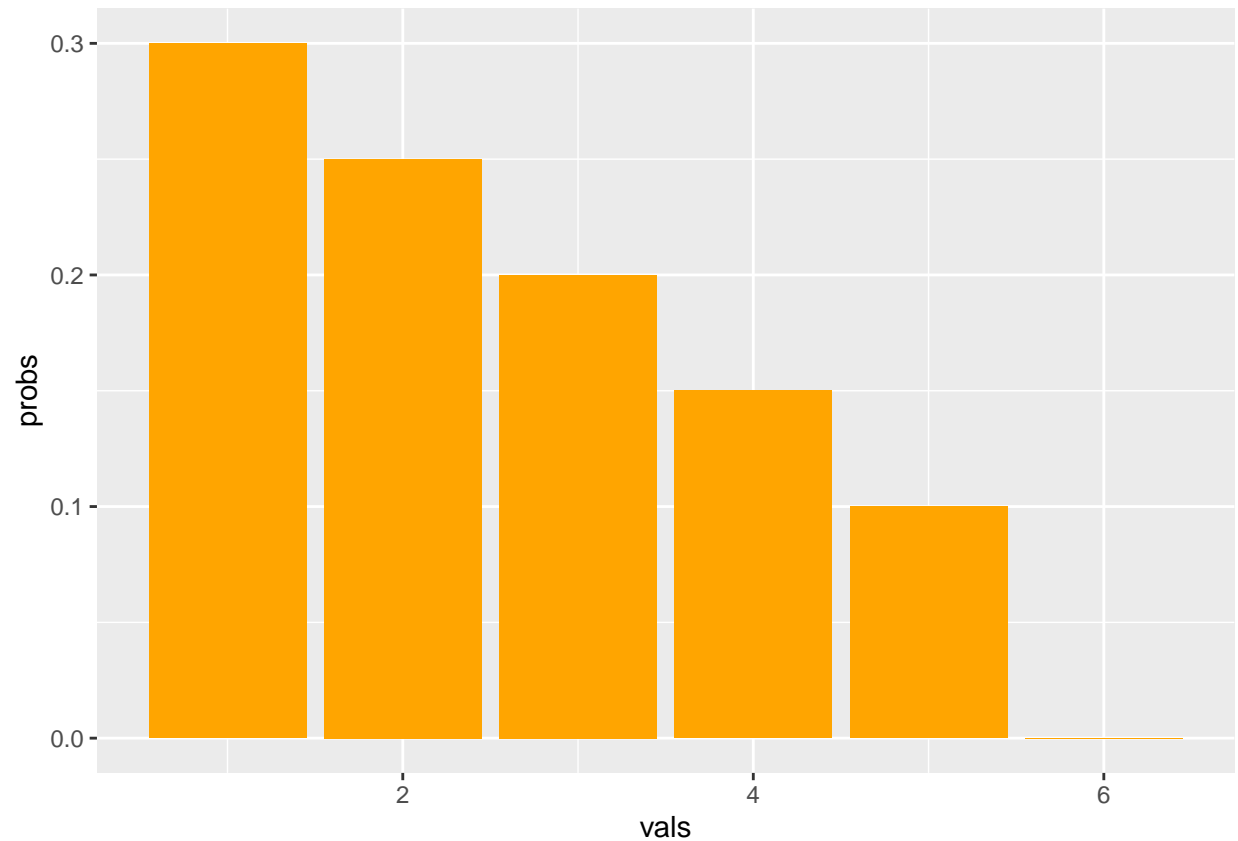
```
get_probs = function(vals) {  
  probs <- vector()  
  
  for (i in seq_along(vals)){  
    if (vals[i] == 6){  
      probs[i] <- 0  
    } else {  
      probs[i] <- (7 - vals[i]) / 20  
    }  
  }  
  return(probs)  
}  
  
vals = c(1, 2, 3, 4, 5, 6)  
probs <- get_probs(vals)
```

Probabilities:

```
## [1] 0.30 0.25 0.20 0.15 0.10 0.00
```

The pmf of the random variable X can be written as such:

$$f(x) = \begin{cases} .30, & x = 1 \\ .25, & x = 2 \\ .20, & x = 3 \\ .15, & x = 4 \\ .10, & x = 5 \\ 0, & otherwise \end{cases}$$



- b) **Determine the cdf of X.**

The cdf of X can be calculated in r as shown below:

```
get_cdf = function(pmf_probs) {
  cdf_probs <- vector()

  for (i in seq_along(pmf_probs)){

    if (i == 1){
      cdf_probs[i] <- pmf_probs[i]
    } else {
      cdf_probs[i] <- pmf_probs[i] + cdf_probs[i - 1]
    }

  }

  return(cdf_probs)
}

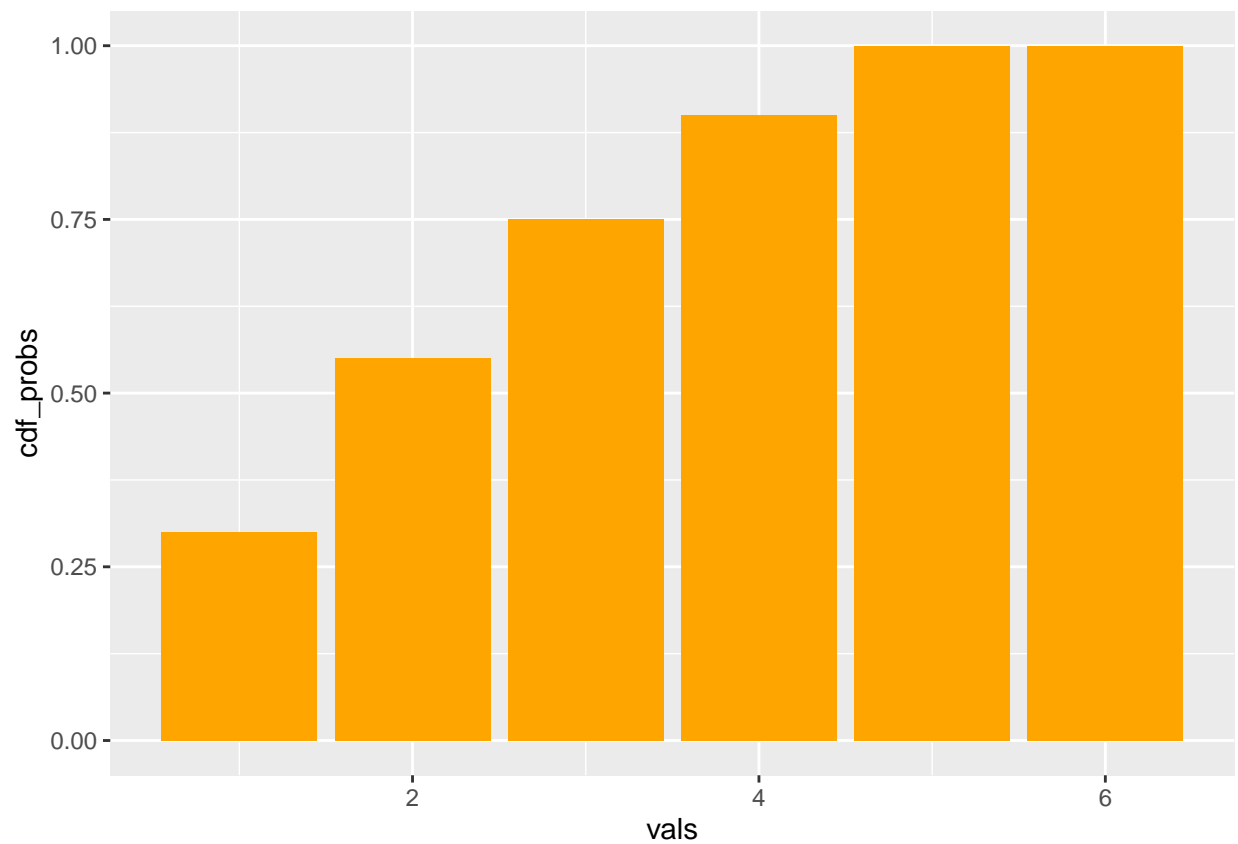
cdf_probs <- get_cdf(probs)

## [1] "Value of X's cdf:"
```

```
## [1] 0.30 0.55 0.75 0.90 1.00 1.00
```

The cdf for  $x$  can be written as shown:

$$F(x) = \begin{cases} 0, & x < 1 \\ .3, & 1 \leq x < 2 \\ .55, & 2 \leq x < 3 \\ .75, & 3 \leq x < 4 \\ .90, & 4 \leq x < 5 \\ 1, & x \geq 5 \end{cases}$$



- c) Determine the expected value of  $X$ .

```
expectation <- sum(vals*probs)
```

```
## [1] "Expected value of X: 2.5"
```

- d) Determine the variance of  $X$ .

```
squared_diff <- (vals - expectation)^2
variance <- sum(squared_diff * probs)
```

```
## [1] "Variance of X: 1.75"
```

- e) Determine the standard deviation of X.

```
stddev <- sqrt(variance)
```

```
## [1] "Standard deviation of X: 1.3228756555323"
```

### Problem 3

- a) Determine the pmf of X.

```
vals = c(1, 2, 5, 10)
probs <- c((4/10), (1/10), (2/10), (3/10))
```

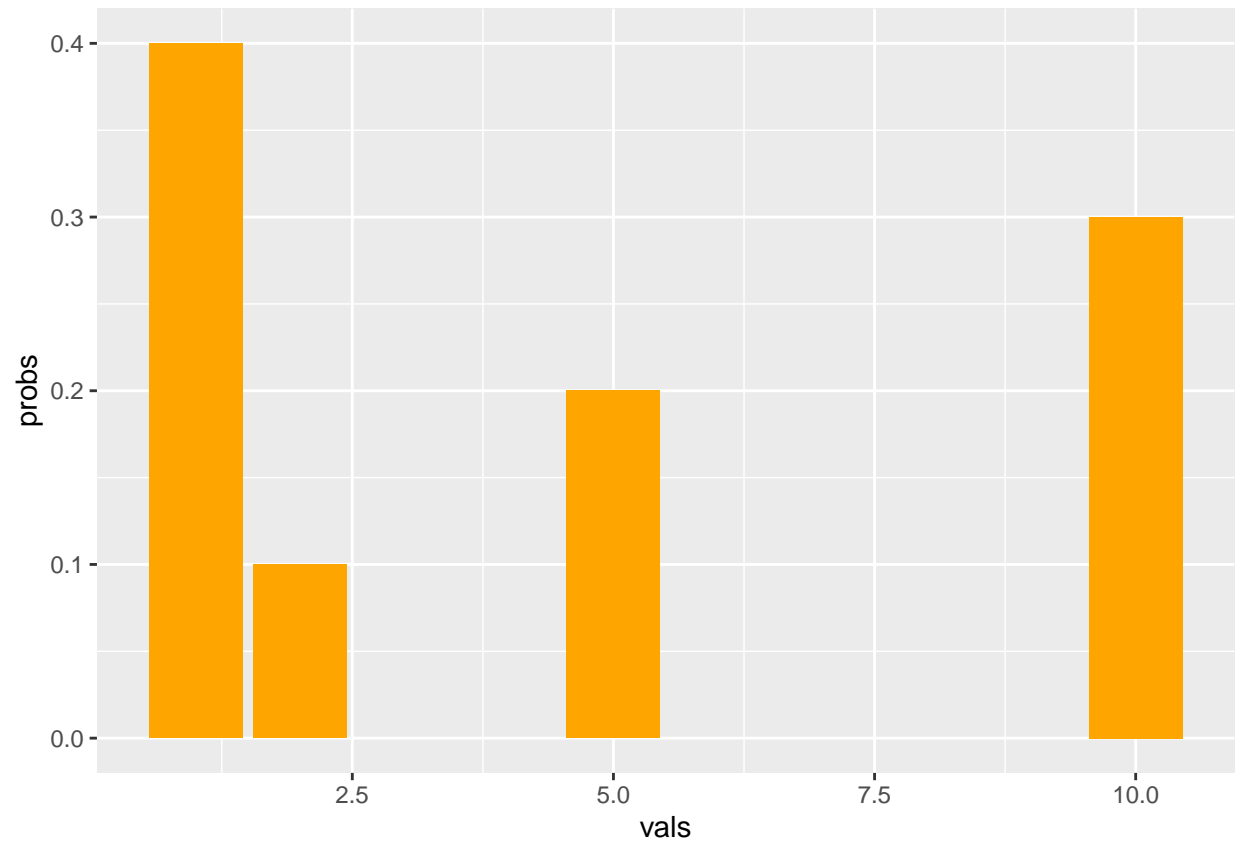
Probabilities:

```
## [1] "Values of X's pmf function:"
```

```
## [1] 0.4 0.1 0.2 0.3
```

X's pmf function can be written as such:

$$f(x) = \begin{cases} .4, & x = 1 \\ .1, & x = 2 \\ .2, & x = 5 \\ .3, & x = 10 \\ 0, & \text{otherwise} \end{cases}$$



- b) **Determine the cdf of X.**

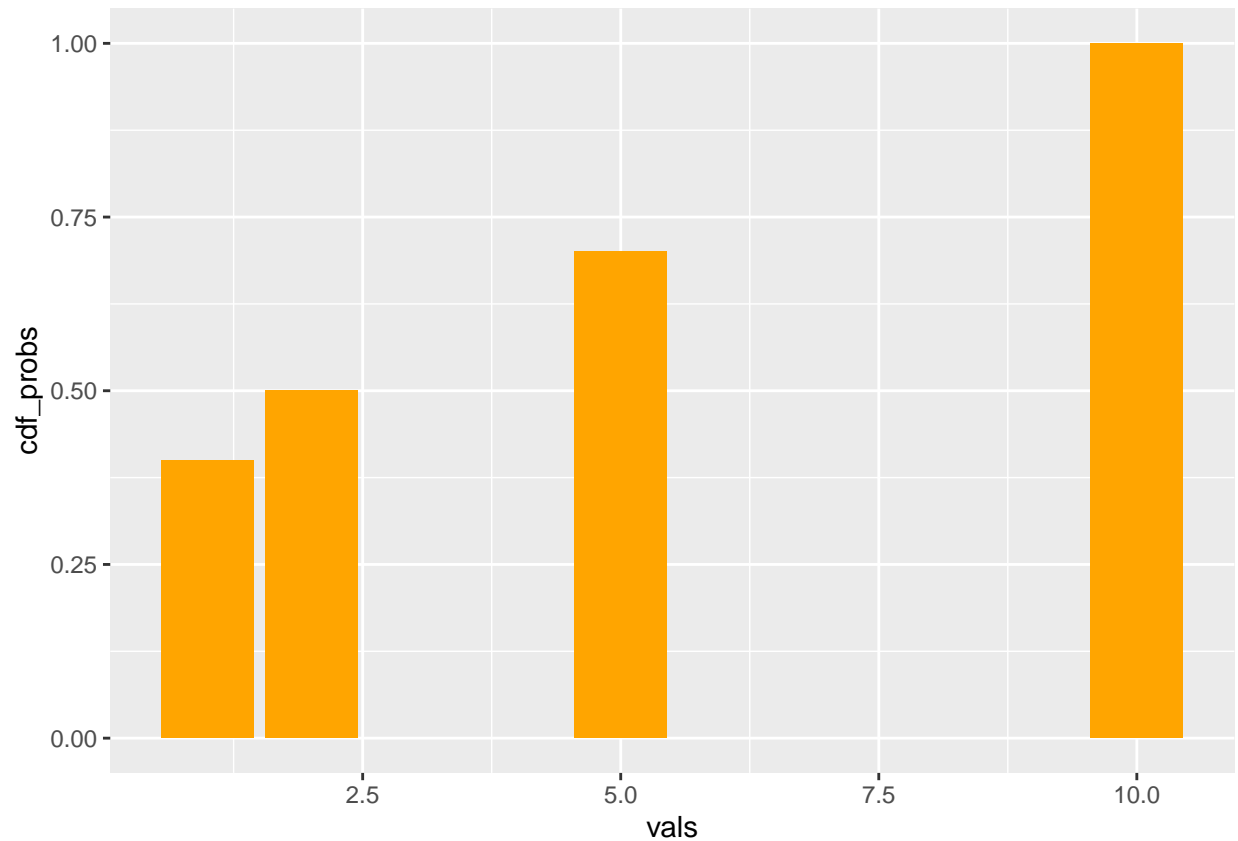
```
cdf_probs = get_cdf(probs)
```

```
## [1] "Values of X's cdf:"
```

```
## [1] 0.4 0.5 0.7 1.0
```

X's cdf can be written as such:

$$F(x) = \begin{cases} 0, & x < 1 \\ .4, & 1 \leq x < 2 \\ .5, & 2 \leq x < 5 \\ .7, & 5 \leq x < 10 \\ 1, & x \geq 10 \end{cases}$$



- c) Determine the expected value of X.

```
expectation <- sum(vals*probs)
```

```
## [1] "Expected value of X: 4.6"
```

- d) Determine the variance of X.

```
squared_diff <- (vals - expectation)^2
variance <- sum(squared_diff * probs)
```

```
## [1] "Variance of X: 14.64"
```

- e) Determine the standard deviation of X.

```
stddev <- sqrt(variance)
```

```
## [1] "Standard deviation of X: 3.8262252939418"
```

## Problem 11

- What is the probability that the hotel will have to turn away statisticians who have reserved rooms?

If we take  $X$  to be the number of statisticians that show up to the hotel, then we can say:

$$X \sim \text{Binomial}(n = 110, p = .96)$$

For the hotel to not turn away any guests, 100 or fewer statisticians need to show up. We can calculate this as such:

$$P(X > 100) = 1 - P(X \leq 100)$$

In R, the calculation can be performed as so:

```
prob <- 1 - pbinom(100, size=110, prob=.96)

## [1] "The probability that the hotel will turn away a guest: 0.987009485801276"
```

## Problem 13

- a) What is the probability that chance variation would result in at least 12 of 38 challenged calls being overturned?

If we take  $X$  to be the number of calls overruled, then we can say:

$$X \sim \text{Binomial}(n = 38, p = .2)$$

The probability that there are 12 or more rulings overturned can be written thusly:

$$P(X \geq 12) = P(X > 11) = 1 - P(X \leq 11)$$

This can be calculated in R as so:

```
prob <- 1 - pbinom(12, size=38, p=.2)

## [1] "The probability that 12 or more calls will be overruled: 0.0287921062236707"
```



- b) Suppose that the William & Mary women's tennis team plays 25 team matches next year and challenges exactly 38 calls in each match. What is the probability that they will play at least one team match in which at least 12 challenged calls are overruled? If we now take X to be the number of matches where 12 or more calls are overruled, we can say:

$$X \sim \text{Binomial}(n = 25, p = .02879)$$

The probability, then, that at least one of 25 matches played throughout a season will have 12 or more calls overturned can be represented as such:

$$P(X \geq 1) = P(X > 0)$$

This can be calculated in R as shown:

```
prob <- pbinom(0, size=35, p=.02879, lower.tail = FALSE)
## [1] "Probability of at least one game having at least 12 calls overturned: 0.640283722614534"
```

**Problem 6:** Let a random variable  $X \sim \text{Bernoulli}(p)$ , compute its population mean  $\mu$ , population variance  $\sigma^2$ , and population standard deviation  $\sigma$ .

The Bernoulli distribution is a special case of the binomial distribution where there is only one experiment performed. This means that there is a single event which has two possible outcomes, 0 or 1, and a probability of success, p. We can therefore calculate the expected value (aka. population mean) of X as so:

$$\begin{aligned}\mu &= 1 * P(X = 1) + 0 * P(X = 0) \\ \mu &= 1 * (p) + 0 * (1 - p) \\ \mu &= p\end{aligned}$$

The variance of X can be calculated as so:

$$\begin{aligned}E(X^2) &= \sum x^2 P(X = x) \\ E(X^2) &= 1^2 * P(X = 1) + 0^2 * P(X = 0) \\ E(X^2) &= 1^2 * p + 0^2 * (1 - p) \\ E(X^2) &= p \\ \sigma^2 &= E(X^2) - E^2(X) \\ \sigma^2 &= p - p^2 \\ \sigma^2 &= p(1 - p)\end{aligned}$$

The standard deviation is then:

$$\sigma = \sqrt{Var(X)}$$

$$\sigma = \sqrt{p(1-p)}$$

**Problem 7:** Let a random variable  $Y \sim \text{Binomial}(n, p)$ , compute its population mean  $\mu$ , population variance  $\sigma^2$ , and population standard deviation  $\sigma$ .

For a variable that follows a Binomial distribution, the probability that that variable takes on a value,  $k$ , can be written as such:

$$P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n$$

This will be useful when calculating the expected value (population mean) and variance of  $Y$ . The expected value of  $Y$  is calculated as shown:

$$\mu = \sum_{k=0}^n k * P(Y = k)$$

$$\mu = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k}$$

$$\mu = \sum_{k=1}^n n \binom{n-1}{k-1} p^k (1-p)^{n-k}$$

$$\mu = np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)}$$

$$\mu = np [p + (1-p)]^{n-1}$$

$$\mu = np$$

Assuming iid., the variance of  $Y$  can then be derived as so:

$$\sigma^2 = Var(X_1 + X_2 + \dots + X_n)$$

$$\sigma^2 = Var(X_1) + Var(X_2) + \dots + Var(X_n)$$

$$\sigma^2 = p(1-p) + p(1-p) + \dots + p(1-p)$$

$$\sigma^2 = np(1 - p)$$

It then follows that the standard deviation,  $\sigma$  is:

$$\begin{aligned}\sigma &= \sqrt{\sigma^2} \\ \sigma &= \sqrt{np(1 - p)}\end{aligned}$$