QBS120_PS1_Correction_Gibran

Gibran Erlangga

9/25/2021

Problems for Rice, Chapter 1

Question 1

my original solution was correct.

Question 2

my original solution was correct.

Question 3

Using the property A of the probability measures, we know that:

$$P(A) = 1 - P(A^C)$$

with $P(A)$ as the probability of one of more defects, and
$P(A^C)$ as the probability of zero defect discovered

With the multiplication principle, we know that the number of ways to select zero defect is the multiplication of # of ways to select m items from $n-k$ non-defects and the # of ways to select no items from k defects. With the total number of outcomes of n choose m, the probability can be represented as

$$P(a \geq 1) = 1 - \frac{\binom{n-k}{m}}{\binom{n}{m}}$$

We need to find m such that $P(a \geq 1) \geq 0.8$. We will perform a computation of $P(a \geq 1)$ in a range of m given n and k values and select m that satisfies above formula.

```
probAtLeastOneDefect = function(n, k ,m) {
  prob = 1 - choose(n-k, m)/choose(n, m)
  return(prob)
}
```

```
m_val = 1:1000
probs = probAtLeastOneDefect(1000, 10, m_val)
(m = which(probs >= 0.8)[1])
```

a) n=1000, k=10

```
## [1] 148
```

```
m_and_probs = data.frame(m_val, probs)
m_and_probs[(m-2):m,]
```

```
##     m_val     probs
## 146   146 0.7952545
## 147   147 0.7976520
## 148   148 0.8000241
```

so, $m = 148$.
b) n=1000, k=100

```
probs = probAtLeastOneDefect(1000, 100, m_val)
(m = which(probs >= 0.8)[1])
```

```
## [1] 16
```

```
m_and_probs = data.frame(m_val, probs)
m_and_probs[(m-2):m,]
```

```
##    m_val     probs
## 14    14 0.7735555
## 15    15 0.7965215
## 16    16 0.8171792
```

so, $m = 16$.

Question 4

The coin flips are all independent, so two coins tossed three times is equal to one coin tossed six times. This is a conditional probability question, with formula as follows

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Since $A \subset B$, we know that $P(A \cap B) = P(A)$, hence

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)}$$

with total possible outcomes of $2^6 = 64$. Since each outcome is equally likely, we can say that $P(A)$ and $P(B)$ as # of possible ways to get A or B divided by total possible outcomes (64). We can also see all the outcomes and count how many of them belongs to A or B:

```
all_outcomes = expand.grid(c1=c("H", "T"),
                           c2=c("H", "T"),
                           c3=c("H", "T"),
                           c4=c("H", "T"),
                           c5=c("H", "T"),
                           c6=c("H", "T"))
tail(all_outcomes)
```

```
##    c1 c2 c3 c4 c5 c6
## 59  H  T  H  T  T  T
## 60  T  T  H  T  T  T
## 61  H  H  T  T  T  T
## 62  T  H  T  T  T  T
## 63  H  T  T  T  T  T
## 64  T  T  T  T  T  T
```

a. What is the probability of two or more heads given that there was at least one head? For easier calculation in the number of outcomes in B, we know that B^C only includes a single outcome with all tails (t, t, t, t, t, t), so

$$P(B) = 1 - P(B^C)$$
$$P(B) = 1 - \frac{1}{64}$$
$$P(B) = \frac{63}{64}$$

Similar process goes for calculating $P(A)$. We know that $P(A^C)$ is essentially the probability of at least five tails. This includes (t, t, t, t, t, t) and outcomes with a single head (which can occupy one of the six available spaces), so there are 6 + 1 = 7 total outcomes in A^C.In other words, $P(A) = 1 - P(A^C) = 1 - 7/64 = 57/64$ Therefore

$$P(A \mid B) = \frac{P(A)}{P(B)} = \frac{57/64}{63/64} = 57/63$$

b. What is the probability of two or more heads given that there was at least one tail? By symmetry, we know that $P(B)$ here has the exact probability as probability of having at least one head in the previous question, so $P(B) = 63/64$. For $A \cap B$, we know that only one event in A (h, h, h, h, h, h) is not also in B so the number of outcomes in $A \cap B$ equals the number in A - 1 or 56, hence $P(A \cap B) = 56/64$. Thus

$$P(A \mid B) = \frac{P(A)}{P(B)} = \frac{56/64}{63/64} = 56/63$$

Question 5

Let A_1 be the event that the first individual has 1 progeny at t = 1, let A_2 be the event that the first individual has 2 progeny at t = 1, let A_3 be the event that the first individual dies with no progeny at t = 1, and let B be the event that all members of the population die at t = 2. We are asked to solve for $P(B)$. We can utilize the law of total probability to define $P(B)$ as follows

$$P(B) = P(A_1 \mid B)P(A_1) + P(A_2 \mid B)P(A_2) + P(A_3 \mid B)P(A_3)$$

Note that this works since A_1, A_2, A_3 are disjoint and $A \cup B \cup C = \omega$. From the question we know that:

$P(A_1) = p$
$P(A_2) = 2p$
$P(A_3) = 1 - P(A_1) - P(A_2) = 1 - 3p$ (since A_3 is the complement of $A_1 \cup A_2$)
$P(B \mid A_3) = 1$ (if the initial individual dies at t=1 then there can be no population at t=2)

To get $P(B \mid A_1)$ and $P(B \mid A_2)$, we know that B requires that all progeny of the first individual die. Since each die with probability $1-3p$ and their deaths are independent, the probability that all die is $1-3p$ if A_1 happens and $(1-3p)^2$ if A_2 happens. Put all values into the original formula, we get:

$$
\begin{aligned}
P(B) &= P(A_1 \mid B)P(A_1) + P(A_2 \mid B)P(A_2) + P(A_3 \mid B)P(A_3) \\
&= P(A_1 \mid B)P(A_1) + P(A_2 \mid B)P(A_2) + P(A_3 \mid B)P(A_3) \\
&= (1-3p)p + (1-3p)^2 2p + (1-3p) \\
&= -15p^2 + 18p^3 + 1
\end{aligned}
$$

Additionally, we can also solve it numerically by computing the value of p such that $P(B) = 0.5$ using R function polyroot().

```
roots = polyroot(c(0.5, 0, -15, 18))
```

```
prob_B = function(x) {
  return(1 - 15*x^2 + 18*x^3)
}

p_val = seq(from=0, to=1, by=0.05)
p_B = sapply(p_val, function(x) {
  return(prob_B(x))
})

names(p_B) = p_val

p_B
```

```
##        0     0.05      0.1     0.15      0.2     0.25      0.3     0.35
##  1.00000  0.96475  0.86800  0.72325  0.54400  0.34375  0.13600 -0.06575
##      0.4     0.45      0.5     0.55      0.6     0.65      0.7     0.75
## -0.24800 -0.39725 -0.50000 -0.54275 -0.51200 -0.39425 -0.17600  0.15625
##      0.8     0.85      0.9     0.95        1
##  0.61600  1.21675  1.97200  2.89525  4.00000
```

Problems for Rice, Chapter 2

Question 1

my original solution was correct.

Question 2

my original solution was correct.

Question 3

Treating each message transmission as a Bernoulli RV with probability of the message succesfully delivered is p. For the majority decoder, we have independent transmissions so the number of successes is a binomial RV with n=3 and p. We will represent the number of successes for the majority decoder as binomial RV.

The probability of successful transmission with the majority decoder, P_D, is:

$$
\begin{aligned}
P_D &= P(X=2) + P(X=3) \\
&= p_X(2) + p_X(3) \\
&= \binom{3}{2} p^2 (1-p)^{3-2} + \binom{3}{3} p^3 (1-p)^{3-3} \\
&= ep^2(1-p) + p^3
\end{aligned}
$$

We want to get the value of p such that $p < P_D$ with both p and P_D bound between 0 and 1.

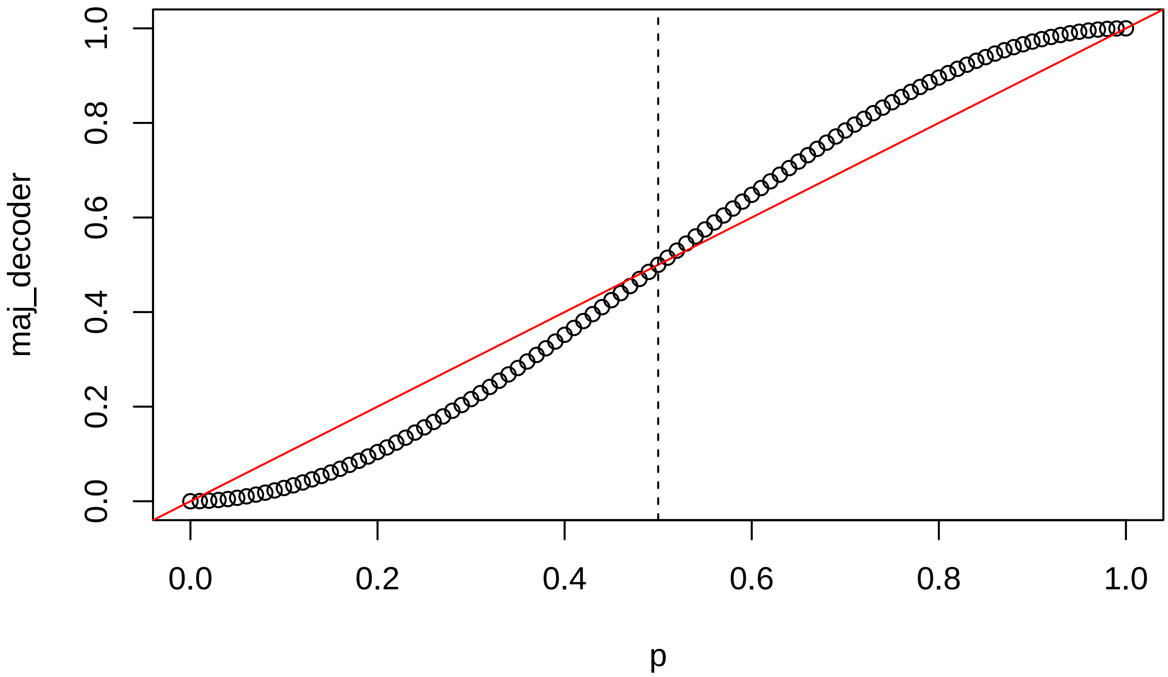
$$
\begin{aligned}
& p < 3p2(1-p) + p^3 \\
& 0 < -2p^3 + 3p^2 - p \\
& 0 < -2p^2 + 3p - 1 \\
& 0 < (-2p+1)(p-1)
\end{aligned}
$$

From the result above, we get the value of p to be either $\frac{1}{2}$ or 1. The possible solution ranges are $p < 0.5$ or $0.5 < p < 1$. Of these results, only p > 0.5 satisfies the inequality. So, the 2 out of 3 majority decoder is superior to direct transmission for p > 0.5.

Validating through visualization in R:

```r
p = seq(from=0, to=1, by=0.01)
maj_decoder = sapply(p, function(x) {
  return(3*x^2*(1-x) + x^3)
})

plot(p, maj_decoder, type="b")
abline(coef = c(0, 1), col="red")
abline(v=0.5, lty="dashed")
```



Question 4

a. Find c. my original solution was correct.

b. Find the CDF. By definition, the CDF for X is:

$$\begin{aligned} F_x(x) &= \int_{-\infty}^{\infty} f_X(u)du \\ &= \int_0^x \frac{3}{20}(u^2+2u)du \\ &= \frac{3}{20}[\frac{u^3}{3}+u^2]_0^x \\ &= \frac{3}{20}(\frac{x^3}{3}+x^2) \end{aligned}$$

CDF for $x < 0$ is 0 and for $x > 2$ is 1. So, can express the CDF as:

c. What is $P(0.1 \leq x \leq 0.5)$? This can be represented as the difference of CDF values at 0.5 and 0.1:

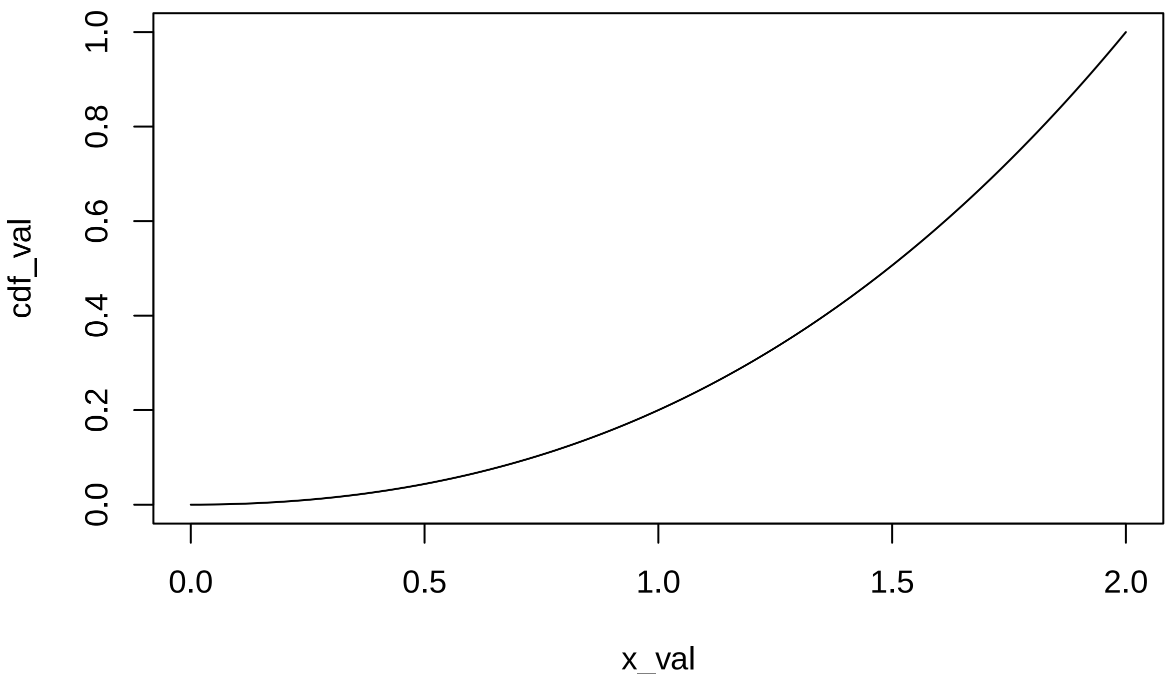
$$\begin{aligned} P(0.1 \leq x \leq 0.5) &= F_X(0.5) - F_X(0.1) \\ &= \frac{3}{20}(\frac{0.5^3}{3}+0.5^2-\frac{0.1^3}{3}-0.1^3) \end{aligned}$$

Computing the result through R:

```
prob_0_1_0_5 = (3/20)*(0.5^3/3 + 0.5^2 - 0.1^3/3 - 0.1^2)
```

```
x_val = seq(from=0, to=2, by=0.01)
cdf_val = sapply(x_val, function(x) {
          return((3/20)*(x^3/3 + x^2))
})

plot(x_val, cdf_val, type="l")
```



d. Plot the CDF.

Question 5

a. What proportion of the population is over 6 ft tall? Question is asking us to find $P(X > 72)$. Using pnorm() we get:

```
1- pnorm(72, mean=60, sd=4)
```

```
## [1] 0.001349898
```

But based on the question, we need to transform the RV into standard normal $N(0, 1)$. We know that normal RV with parameters μ and std, $(X-)/std$ is standard normal. Hence

$$
\begin{aligned}
P_X(X > 72) &= P_X((X - 60)/4 > (72 - 60)/4) \\
&= P_Z(Z > 3) \\
&= 1 - \Phi(3)
\end{aligned}
$$

Then we can run the formula in R:

```
1 - pnorm(3)
```

```
## [1] 0.001349898
```

b. What is the distribution of heights if they are expressed in centimeters? Do scaling transformation to centimeters:

$$N(60 * 2.54, 2.54^2 * 4^2) = N(152.4, 103.2256)$$

c. In meters? Do scaling transformation to centimeters:

$$N(60 * 0.0254, 0.0254^2 * 4^2) = N(1.524, 0.01032256)$$