Stats for DS HW 3

Matthew DeSantis

2022-09-20

2.3.15

(a)

```
48/choose(52,5)
```

```
## [1] 1.846893e-05
```

(b)

```
choose(4,2)*choose(4,2)*44/choose(52,5)
```

```
## [1] 0.0006094746
```

(c)

```
choose(4,3)*choose(12,2)*4*4/choose(52,5)
```

```
## [1] 0.001625265
```

2.4.6

(a)

E1 and not E2. This is equal to P(E1) - P(E1 and E2).

```
0.95 - 0.88
```

```
## [1] 0.07
```

(b)

E2 and not E1. This is equal to P(E2) - P(E2 and E1).

```
0.92 - 0.88
```

```
## [1] 0.04
```

(c)

(E1 and not E2) or (E2 and not E1). This is equal to the sum of the answers from (a) and (b).

```
(0.95-0.88)+(0.92-0.88)
```

```
## [1] 0.11
```

(d)

E1 or E2. This is equal to P(E1) + P(E2) - P(E1 and E2).

```
0.95+0.92-0.88
```

```
## [1] 0.99
```

2.5.12

(a)

E1 = Aircraft is discovered = 0.7. E2 = Aircraft has locator. P(E2 | E1) = 0.6. P(E2 | not E1) = 0.1. P(E2 and not E1) = P(E2 | not E1)*P(not E1) =

```r
0.1*0.3
```

```
## [1] 0.03
```

(b)

P(E2) = P(E1)*P(E2 | E1) + P(not E1)*P(E2 | not E1) =

```r
(0.7*0.6) + (0.3*0.1)
```

```
## [1] 0.45
```

(C)

P(not E1 | E2) = (P(not E1)*P(E2 | not E1))/(P(not E1)*P(E2 | not E1) + P(E1)*P(E2 | E1)) =

```r
(0.3*0.1)/((0.3*0.1)+(0.7*0.6))
```

```
## [1] 0.06666667
```

2.6.7

(a)

Table 1: Sports Preferences by Gender

	Football	Basketball	Track	Total
Male	0.3	0.22	0.13	0.65
Female	0.0	0.28	0.07	0.35
Total	0.3	0.50	0.20	1.00

(b)

P(Female | Basketball) = P(Female and Basketball)/P(Basketball) =

```
0.28/0.5
```

```
## [1] 0.56
```

(c)

No. Because P(F | B) != P(F), (0.56 != 0.5), by definition, they are not independent.

3.2.4

```
total = choose(10,3)
PMF = data.frame(zero = c(choose(7,3)/total),
                 one = c(choose(3,1)*choose(7,2)/total),
                 two = c(choose(3,2)*choose(7,1)/total),
                 three = c(choose(3,3)/total)
                )
rownames(PMF) = c("P(X=x)")
kable(PMF, caption = "PMF")
```

Table 2: PMF

	zero	one	two	three
P(X=x)	0.2916667	0.525	0.175	0.0083333

```
CDF = data.frame(zero = c(choose(7,3)/total),
                 one = c(choose(3,1)*choose(7,2)/total+choose(7,3)/total),
                 two = c(choose(3,2)*choose(7,1)/total+
                           choose(3,1)*choose(7,2)/total+choose(7,3)/total),
                 three = c(choose(3,3)/total+choose(3,2)*choose(7,1)/total+
                             choose(3,1)*choose(7,2)/total+choose(7,3)/total)
                )
rownames(CDF) = c("P(X=x)")
kable(CDF, caption = "CDF")
```

Table 3: CDF

	zero	one	two	three
P(X=x)	0.2916667	0.8166667	0.9916667	1

3.2.8

(a)

```
checkout_CDF = function(x) {
  if (x<=0){
    return (0)
  }

  if (0<x && x<=2) {
    return ((x**2)/4)
  }

  else {
    return (1)
  }
}

checkout_CDF(1)-checkout_CDF(0.5)
```

```
## [1] 0.1875
```

(b)

The PDF can be obtained from the CDF by integrating with respect to x. Doing so in this case results in f(x) = x/2 for x between 0 and 2, 0 otherwise.

(c)

To derive these, substitute y/60 for x. Therefore, the CDF is ((y/60)ˆ2)/4 is: CDF of Y = (yˆ2)/14400 for y between 0 and 120, 0 otherwise. Taking the derivative, we get PDF of Y = y/7200 for y between 0 and 2, 0 for x<=0, and 1 for x>120.