Midterm 2 - Solutions

Statistics - NYU, Summer 2016
Ercan Karadas

[1] M: Men
W: Women
G: Graduate training
UG: Undergraduate training
HS: High school training

Then let us write what we are given in the problem: $P(M) = .8$, $P(W) = .2$, $P(G|M) = .1$, $P(HS \mid M) = .6$, $P(UG|M) = .3$, $P(G|W) = .15$, $P(UG|W) = .4$, $P(HS|W) = .45$

a) $P(W \cap HS) = P(HS \mid W)P(W) = (.45)(.2) = .09$

b) $P(G) = P(M \cap G) + P(W \cap G) = P(G|M)P(M) + P(G|W)P(W) = (.10)(.8) + (.15)(.2) = .11$, hence $P(\overline{G}) = 1 - 0.11 = .89$

c) $P(W|G) = P(W \cap G)/P(G) = P(G|W)P(W)/P(G) = (.15)(0.2)/.11 = .27$

d) $P(M|\bar{G}) = P(M \cap \bar{G})/P(\bar{G}) = \frac{[P(M) - P(M \cap G)]}{1 - P(G)} = \frac{(.8 - (.10)(.8))}{.89} = .81$

[2] a) $X \in \{0, 1, 2, 3\}$.

b)

$$
\begin{aligned}
P(X = 0) &= (0.1)(0.1)(0.2) = 0.002 \\
P(X = 1) &= 2(0.9)(0.1)(0.2) + (0.1)(0.1)(0.8) = .044 \\
P(X = 2) &= (0.9)(0.9)(0.2) + 2(0.1)(0.9)(0.8) = 0.306 \\
P(X = 3) &= (0.9)(0.9)(0.8) = 0.648
\end{aligned}
$$

Then the probability distribution table is given by:

X	$P(X)$
0	0.002
1	0.044
2	0.306
3	0.648

And

$$
P(X = 0 \mid X \leq 2) = \frac{0.002}{0.001 + 0.044 + 0.306} = \frac{0.002}{0.0351} = 0.00568
$$

c) $E(X) = 2.6$, so expected revenue is $2.6 \times 650 = 1690$.

d) Let Y be r.v. to denote the number of days with at most two vehicles in use in a given day. Then, success probability is $p = P(X \leq 2) = 1 - 0.648 = 0.352$, and using binomial model with parameters $p = 0.352$ and $n = 90$, i.e. $B(0.352, 90)$, we can find $P(Y \leq 30)$ exactly. However, computing this probability is very hard so let us check whether we can use normal approximation to binomial. Since $np(1-p) = 90(.352)(.648) = 20.5 > 5$, we can use normal approximation here, namely we can consider as if Y is normally distributed with a mean $90(0.352) = 31.68$ and variance $90(0.352)(1-0.352) = 20.5$, that is $Y \sim N(31.68, 20.5)$. Therefore,

$$P(Y < 30) = P(Z < \frac{30 - 31.68}{\sqrt{20.5}}) = P(Z \leq -0.37) \approx 0.355.$$

[3] a) $X \sim N(20, 4^2)$
$P(15 < X < 25) = P(\frac{15-20}{4} < Z < \frac{25-20}{4}) = P(-1.25 < Z < 1.25) = F(1.25) - F(-1.25) = (0.8944) - (1 - 0.8944) = 0.7888$

b) Find an x value such that $P(X < x) = 0.90$.
$P(\frac{X-20}{4} < \frac{x-20}{4}) = P(Z < \frac{x-20}{4}) = 0.90$, then desired can be found by solving $F(\frac{x-20}{4}) = 0.90$ or equivalently $\frac{x-20}{4} = 1.28$. Doing so yields $x = 25.12$

c) The interval should be centered at the mean value, and therefore 20% percent should remain in each tail. In order to find the upper bound of the interval we need to find an x value such that $P(X < x) = 0.80$. Similar to part (b), this value is given by the equation $\frac{x-20}{4} = 0.84$ which yields $x = 23.36$. Since the interval is symmetric around the mean 20, the lower bound is equal to 16.64. Therefore shortest range is 23.36-16.64=6.72

d) $P(X > 30) = P(Z > \frac{30-20}{4}) = P(Z > 2.5) = 1 - F(2.5) = 1 - 0.9938 = 0.0062$
This is the probability of getting a free pizza for a random chosen order.

Now, let Y be the number of free pizzas that a student who orders pizza in five consecutive evenings gets, then Y is Binomial with $p = 0.0062$ and $n = 5$, and hence

$$\begin{aligned} P(Y \geq 1) &= 1 - P(Y = 0) \\ &= 1 - \left[\binom{5}{0}(0.0062)^0(1 - 0.0062)^5\right] \\ &= 1 - (0.969)^5 = 0.031. \end{aligned}$$

[4]

a)

Y	0	1	2
$P(y)$	.28	.26	.46

b) $\mu_y = \sum_{y\in\{0,1,2\}} yP(Y=y) = 0(0.28) + 1(0.26) + 2(0.46) = 1.18$

$$\begin{aligned}\sigma_y^2 &= \sum_{y\in\{0,1,2\}} (y-\mu_y)^2 P(Y=y) \\ &= (0-1.18)^2(0.28) + (1-1.18)^2(0.26) + (2-1.18)^2(0.46) \\ &= 0.7076\end{aligned}$$

Therefore, $\sigma_y = \sqrt{0.7076} = 0.841$

c) First construct the prob. dist. table for the r.v. $Y|X=3$

$Y\|X=3$	0	1	2
$P(y\|X=3)$	0.06/0.45	0.11/0.45	0.28/0.45

Then

$$E(Y|X=3) = 0(0.06/0.45) + 1(0.11/0.45) + 2(0.28/0.45) = 1.488$$

d)

T	0	3	13
$P(t)$	.28	.26	.46

and hence $E(T) = 6.3$

e) There are two combinations of (X,Y) that gives $X+Y=4$:

$$(X=2, Y=2), \text{ and } (X=3, Y=1),$$

so $P(X+Y=4) = 0.16 + 0.11 = 0.27$.

To calculate $E(W|X+Y=4)$, first compute $W = 5X + 2Y$ for each of these two pairs, and then sum up over these values after weighting with the corresponding conditional probabilities:

$$\begin{aligned}E(W|X+Y=4) &= [5(2)+2(2)]\frac{0.16}{0.27} + [5(3)+2(1)]\frac{0.11}{0.27} \\ &= 15.22\end{aligned}$$