ST-540 Assignment-2

Tilekbek Zhoroev

Problem 1 X_1 and X_2 have joint PMF

x_1	x_2	$Prob(X_1 = x_1; X_2 = x_2)$
0	0	0.15
1	0	0.15
2	0	0.15
0	1	0.15
1	1	0.20
2	1	0.20

(a) Compute the marginal distribution of X_1.
(b) Compute the marginal distribution of X_2.
(c) Compute the conditional distribution of $X_1|X_2$.
(d) Compute the conditional distribution of $X_2|X_1$.
(e) Are X1 and X2 independent? Justify your answer.

Solution

(a)

$$P(X_1 = 0) = P(X_1 = 0, X_2 = 0) + P(X_1 = 0, X_2 = 1) = .15 + .15 = .3$$

$$P(X_1 = 1) = P(X_1 = 1, X_2 = 0) + P(X_1 = 1, X_2 = 1) = .15 + .2 = .35$$

$$P(X_1 = 2) = P(X_1 = 2, X_2 = 0) + P(X_1 = 2, X_2 = 1) = .15 + .3 = .35$$

(b)

$$P(X_2 = 0) = P(X_1 = 0, X_2 = 0) + P(X_1 = 1, X_2 = 0) + P(X_1 = 2, X_2 = 0) = .15 + .15 + .15 = .45$$

$$P(X_2 = 1) = P(X_1 = 0, X_2 = 1) + P(X_1 = 1, X_2 = 1) + P(X_1 = 2, X_2 = 1) = .15 + .2 + .2 = .55$$

(c)

$$P(X_1 = 0|X_2 = 0) = \frac{P(X_1=0, X_2=0)}{P(X_2=0)} = \frac{.15}{.45} = \frac{1}{3}$$

$$P(X_1 = 1|X_2 = 0) = \frac{P(X_1=1, X_2=0)}{P(X_2=0)} = \frac{.15}{.45} = \frac{1}{3}$$

$$P(X_1 = 2|X_2 = 0) = \frac{P(X_1=2, X_2=0)}{P(X_2=0)} = \frac{.15}{.45} = \frac{1}{3}$$

$$P(X_1 = 0|X_2 = 1) = \frac{P(X_1=0, X_2=1)}{P(X_2=1)} = \frac{.15}{.55} = \frac{3}{11}$$

$$P(X_1 = 1|X_2 = 1) = \frac{P(X_1=1, X_2=1)}{P(X_2=1)} = \frac{.2}{.55} = \frac{4}{11}$$

$$P(X_1 = 2|X_2 = 1) = \frac{P(X_1=2, X_2=1)}{P(X_2=1)} = \frac{.2}{.55} = \frac{4}{11}$$

(d)

$$P(X_2 = 0|X_1 = 0) = \frac{P(X_1=0, X_2=0)}{P(X_1=0)} = \frac{.15}{.3} = \frac{1}{2} = .5$$

$P(X_2 = 1|X_1 = 0) = \frac{P(X_1=0,X_2=1)}{P(X_1=0)} = \frac{.15}{.3} = \frac{1}{2} = .5$

$P(X_2 = 0|X_1 = 1) = \frac{P(X_1=1,X_2=0)}{P(X_1=1)} = \frac{.15}{.35} = \frac{3}{7} = .5$

$P(X_2 = 1|X_1 = 1) = \frac{P(X_1=1,X_2=1)}{P(X_1=1)} = \frac{.2}{.35} = \frac{4}{7}$

$P(X_2 = 0|X_1 = 2) = \frac{P(X_1=2,X_2=0)}{P(X_1=2)} = \frac{.15}{.35} = \frac{3}{7}$

$P(X_2 = 1|X_1 = 2) = \frac{P(X_1=2,X_2=1)}{P(X_1=2)} = \frac{.2}{.35} = \frac{4}{7}$

(e) Since
$$P(X_1 = 1, X_2 = 1) = 0.2 \neq 0.1925 = .35 \times .55 = P(X_1 = 1) \times P(X_2 = 1),$$
X_1 and X_2 are dependent.

Problem 2

Assume $(X_1; X_2)$ have bivariate PDF

$$f(x_1; x_2) = \frac{1}{2\pi}(1 + x_1^2 + x_2^2)^{-\frac{3}{2}}$$

(a) Plot the conditional distribution of $X_1|X_2 = x_2$ for $x_2 \in \{-3, -2, -1, 0, 1, 2, 3\}$ (preferably on the same plot).

(b) Do X_1 and X_2 appear to be correlated? Justify your answer.

(c) Do X_1 and X_2 appear to be independent? Justify your answer

Solution

(a) Here we used conditional probability

$$P_{X_1|X_2}(x_1|x_2) = \frac{f(x_1, x_2)}{P_{X_2}(x_2)}$$

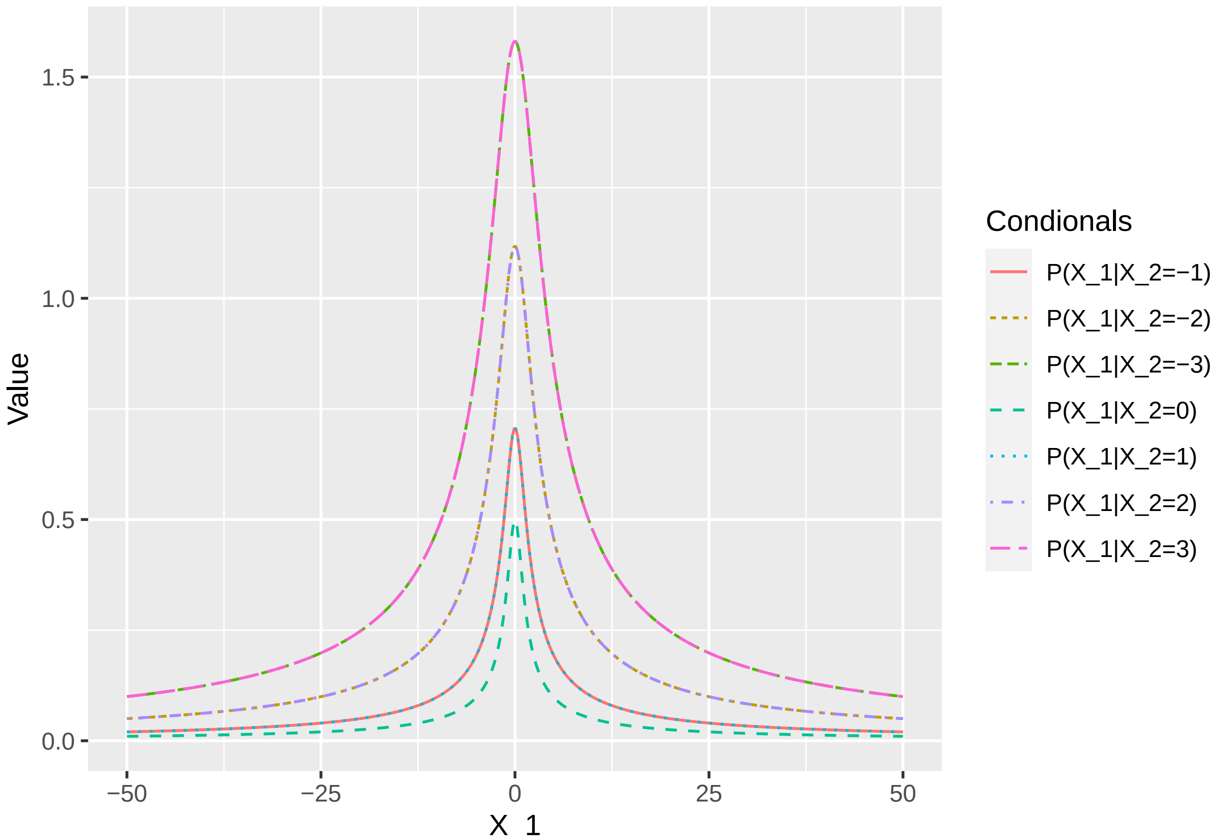
Then the marginal probability is
$$P_{X_2}(x_2) = \frac{1}{(1 + x_2^2)\pi}$$

Then
$$P_{X_1|X_2}(x_1|x_2) = \frac{1}{2}(1 + x_1^2 + x_2^2)^{-\frac{3}{2}}((1 + x_2^2))$$

```r
x1<-seq(-50,50, length.out=500) # let us take x1 points between -50 and 50
x2<-c(-3,-2,-1,0,1,2,3) # x2 are given to us

con_dist<-sapply(x2,pdf_func) # apply function to all points in x2 by for loop
df<-data.frame(x1, con_dist) #define data frame
colnames(df)<-c("X_1","P(X_1|X_2=-3)","P(X_1|X_2=-2)","P(X_1|X_2=-1)",
                "P(X_1|X_2=0)","P(X_1|X_2=1)","P(X_1|X_2=2)","P(X_1|X_2=3)")
df<- df%>%
pivot_longer(cols = "P(X_1|X_2=-3)":"P(X_1|X_2=3)", names_to = "Condionals",
             values_to = "Value") #I used this to create graphs in one figure

## use ggplot
ggplot(df, aes(x=X_1, y=Value, color=Condionals, linetype = Condionals  ))+
        geom_line(aes(color = Condionals))
```



(b) Using $P_{X_1}(x_1) = \frac{1}{(1+x_1^2)\pi}$ we calculate the expectation as

$$E[X_1] = \int_{-\infty}^{\infty} \frac{x_1}{(1+x_1^2)\pi}\, dx_1 = 0$$

and

$$E[X_2] = \int_{-\infty}^{\infty} \frac{x_2}{(1+x_2^2)\pi}\, dx_1 = 0$$

and

$$E[X_1 X_2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{x_1 x_2}{2\pi(1+x_1^2+x_2)}\, dx_1 = 0$$

Thus, the covariance is zero implies that X_1 and X_2 are uncorrelated.

(c) There are not independent, since $\frac{1}{2\pi}(1+x_1^2+x_2^2)^{-\frac{3}{2}} = P(X_1, X_2) \neq P(X_1)P(X_2) = \frac{1}{(1+x_1^2)\pi} \cdot \frac{1}{(1+x_2^2)\pi}$.

Problem 3

For this problem pretend we are dealing with a language with a six-word dictionary

$$\{\text{fun, sun, sit, sat, fan, for}\}.$$

An extensive study of literature written in this language reveals that all words are equally likely except that "for" is α times as likely as the other words. Further study reveals that:

i. Each keystroke is an error with probability θ.

ii. All letters are equally likely to produce errors.

iii. Given that a letter is typed incorrectly it is equally likely to be any other letter.

iv. Errors are independent across letters.

For example, the probability of correctly typing "fun" (or any other word) is $(1-\theta)^3$, the probability of typing "pun" or "fon" when intending to type is "fun" is $\theta(1-\theta)^2$, and the probability of typing "foo" or "nnn" when intending to type "fun" is $\theta^2(1-\theta)$. Use Bayes' rule to develop a simple spell checker for this language. For each of the typed words "sun", "the", "foo", give the probability that each word in the dictionary was the intended word. Perform this for the parameters below:

(a) $\alpha = 2$ and $\theta = 0.1$

(b) $\alpha = 50$ and $\theta = 0.1$

(c) $\alpha = 2$ and $\theta = 0.95$

Comment on the changes you observe in these three cases

Solution (a)

```
likelihood = function(theta, v){
  theta^v*(1-theta)^(3-v)
}

alpha= 2; theta=.1
v_sun=c(1,0,2,2,2,3)
v_the=c(3,3,3,3,3,3)
v_foo=c(2,3,3,3,2,1)
data=data.frame(prior=round(ifelse(c("fun","sun","sit","sat","fan","for")=="for",
                                   alpha/(5+alpha),1/(5+alpha)),3))
row.names(data) = c("fun","sun","sit","sat","fan","for")
data<-data%>%
  mutate(sun_likelihood = round(likelihood(theta,v_sun),3))%>%
  mutate(sun_posterior = round(sun_likelihood*prior/sum(sun_likelihood*prior),3))%>%
  mutate(the_likelihood = round(likelihood(theta,v_the),3))%>%
  mutate(the_posterior= round(the_likelihood*prior/sum(the_likelihood*prior),3))%>%
  mutate(foo_likelihood = round(likelihood(theta,v_foo),3))%>%
  mutate(foo_posterior= round(foo_likelihood*prior/sum(foo_likelihood*prior),3))
data[,c(1,3,5,7)]
```

```
##     prior sun_posterior the_posterior foo_posterior
## fun 0.143         0.097         0.143         0.049
## sun 0.143         0.869         0.143         0.005
## sit 0.143         0.011         0.143         0.005
## sat 0.143         0.011         0.143         0.005
## fan 0.143         0.011         0.143         0.049
## for 0.286         0.002         0.286         0.885
```

(b)

```
likelihood = function(theta, v){
  theta^v*(1-theta)^(3-v)
}

alpha= 50; theta=.1
v_sun=c(1,0,2,2,2,3)
v_the=c(3,3,3,3,3,3)
v_foo=c(2,3,3,3,2,1)
data=data.frame(prior=round(ifelse(c("fun","sun","sit","sat","fan","for")=="for",
                                   alpha/(5+alpha),1/(5+alpha)),3))
row.names(data) = c("fun","sun","sit","sat","fan","for")
data<-data%>%
```

```r
  mutate(sun_likelihood = round(likelihood(theta,v_sun),3))%>%
  mutate(sun_posterior = round(sun_likelihood*prior/sum(sun_likelihood*prior),3))%>%
  mutate(the_likelihood = round(likelihood(theta,v_the),3))%>%
  mutate(the_posterior= round(the_likelihood*prior/sum(the_likelihood*prior),3))%>%
  mutate(foo_likelihood = round(likelihood(theta,v_foo),3))%>%
  mutate(foo_posterior= round(foo_likelihood*prior/sum(foo_likelihood*prior),3))
data[,c(1,3,5,7)]
```

```
##     prior sun_posterior the_posterior foo_posterior
## fun 0.018         0.091         0.018         0.002
## sun 0.018         0.821         0.018         0.000
## sit 0.018         0.010         0.018         0.000
## sat 0.018         0.010         0.018         0.000
## fan 0.018         0.010         0.018         0.002
## for 0.909         0.057         0.910         0.995
```

Here we notice that when we increase α the posterior probability of error on α also increased since the priors are changed. The other probabilities are relatively decreased.

(c)

```r
likelihood = function(theta, v){
  theta^v*(1-theta)^(3-v)
}

alpha= 2; theta=.95
v_sun=c(1,0,2,2,2,3)
v_the=c(3,3,3,3,3,3)
v_foo=c(2,3,3,3,2,1)
data=data.frame(prior=round(ifelse(c("fun","sun","sit","sat","fan","for")=="for",
                                   alpha/(5+alpha),1/(5+alpha)),3))
row.names(data) = c("fun","sun","sit","sat","fan","for")
data<-data%>%
  mutate(sun_likelihood = round(likelihood(theta,v_sun),3))%>%
  mutate(sun_posterior = round(sun_likelihood*prior/sum(sun_likelihood*prior),3))%>%
  mutate(the_likelihood = round(likelihood(theta,v_the),3))%>%
  mutate(the_posterior= round(the_likelihood*prior/sum(the_likelihood*prior),3))%>%
  mutate(foo_likelihood = round(likelihood(theta,v_foo),3))%>%
  mutate(foo_posterior= round(foo_likelihood*prior/sum(foo_likelihood*prior),3))
data[,c(1,3,5,7)]
```

```
##     prior sun_posterior the_posterior foo_posterior
## fun 0.143         0.001         0.143         0.017
## sun 0.143         0.000         0.143         0.322
## sit 0.143         0.024         0.143         0.322
## sat 0.143         0.024         0.143         0.322
## fan 0.143         0.024         0.143         0.017
## for 0.286         0.926         0.286         0.002
```

When we increase θ, if we have all errors (all letters are different) then posterior probability become very large, on the other hand partially match or correct decreases the posterior probability. To sum up the distribution of probabilities are more diverse than case (a).

Problem 4

If 70% of a population is vaccinated, and the hospitalization rate is 5 times higher for an unvaccinated person than a vaccinated person, what is the probability that a person is vaccinated given they are hospitalized?

Solution Let θ = vaccinated and Y = hospitalization. We have given that $P(\theta) = .7$ and $P(Y|\theta) = 1/6$. We need to find $P(\theta|Y)$. Using Bayes Theorem

$$P(\theta|Y) = \frac{P(Y|\theta)P(\theta)}{P(Y|\theta)P(\theta) + P(Y|\theta')P(\theta')} = \frac{1/6 * .7}{1/6 * .7 + 5/6 * .3} \approx 0.3182$$