STAT 445

Theoretical Statistics I

Fall Semester 2017

Final Exam

Name: Solution

- You have 180 min to complete this exam
- Justify your answers
- Evaluate expressions as much as you can
- There is one blank page in the back in case you need more space for any of your answers.

Problem	1	2	3	4	5	6	7	8	9	Total
Missed Score										
out of	17	23	5	8	6	8	10	15	8	100

Good luck!

1. The Naval Facility in Keflavik, Iceland, was tracking a Soviet submarine south through the Norwegian Sea for several days, then lost it[1]. Your squadron has been tasked to search for it. The submarine could be in one of three areas:

   - A1) west of Iceland, transiting the Greenland–Iceland gap
   - A2) east of Iceland, transiting the Iceland–UK gap
   - A3) it might have changed direction to take up a patrol station in the western portion of the Norwegian Sea.

   You will send planes to search for the submarine in all three areas. Based on historical patterns, you estimate the probability that the submarine went to areas A1, A2, and A3, as 10%, 75%, and 15%, respectively. However, a submarine is easier to detect in some areas than others. Knowing the area quite well, you know that if the submarine is in area A1 the probability that it is detected is 60%. If the submarine is in area A2 the probability of detection is rather high, 80%, but if it is in area A3 the detection probability is only 40%.

   (a) (5 points) You conduct a search of all three locations. What is the probability that you do not detect the submarine in this first search?

Define events:

A1 : submarine is in area A1

A2 : — " — " — A2

A3 : — " — " — A3

D : submarine is detected

Given: $P(A1) = 0.1$, $P(A2) = 0.75$, $P(A3) = 0.15$

$P(D \mid A1) = 0.6$, $P(D \mid A2) = 0.8$, $P(D \mid A3) = 0.4$

Want: $P(D^c) = 1 - P(D)$

$$= 1 - \left[P(D \mid A1)P(A1) + P(D \mid A2)P(A2) + P(D \mid A3)P(A3)\right]$$

$$= 1 - \left[0.6 \cdot 0.1 + 0.8 \cdot 0.75 + 0.4 \cdot 0.15\right]$$

$$= 1 - 0.72 = \underline{0.28}$$

[1] A typical scenario in the Anti-Submarine Warfare situations in the 1980s.

(b) (4 points) If your first search did not detect the submarine what are the probabilities that the submarine is in each location, A1, A2, and A3?

$$P(A1|D^c) = \frac{P(D^c|A1)P(A1)}{P(D^c)} = \frac{0.4 \cdot 0.1}{0.28} = 0.1429$$

$$P(A2|D^c) = \frac{0.2 \cdot 0.75}{0.28} = 0.5357$$

Most likely area is still A2

$$P(A3|D^c) = \frac{0.6 \cdot 0.15}{0.28} = 0.3214$$

(c) (4 points) Assuming that your first search was unsuccessful, you conduct a second search of all three locations. What is the probability that you do not detect the submarine in this second search either?

Note that detection probabilities are independent of how many prior search have been conducted.

Define: D_1: detected in first round

D_2: — " — second round

$$P(D_2|Ak, D_1^c) = P(D|Ak) \text{ for } k=1,2,3$$

So, $$P(D_2^c|D_1^c) = P(D_2^c|A1,D_1^c)P(A1|D_1^c) + P(D_2^c|A2,D_1^c)P(A2|D_1^c) + P(D_2^c|A3,D_1^c)P(A3|D_1^c)$$

$$= 0.4 \cdot 0.1429 + 0.2 \cdot 0.5357 + 0.6 \cdot 0.3214$$

$$= 0.357$$

(d) (4 points) At this point (two failed searches), in what location is the submarine most likely to be located?

$$P(A1|D_1^c,D_2^c) = \frac{P(D_2^c|A1,D_1^c)P(A1|D_1^c)}{P(A1|D_1^c)} = \frac{0.4 \cdot 0.1429}{0.357} = 0.16$$

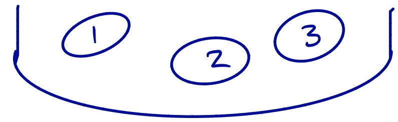
$$P(A2|D_1^c,D_2^c) = \frac{0.2 \cdot 0.5357}{0.357} = 0.30$$

$$P(A3|D_1^c,D_2^c) = \frac{0.6 \cdot 0.3214}{0.357} = 0.54$$

Most likely area is now A3, where detection probability is lowest.

2. Consider a sample of size two drawn without replacement from an urn containing three balls, numbered 1, 2 and 3. Let X be the number on the first ball drawn and Y the larger of the two numbers drawn.

(a) (6 points) Find the joint probability mass function (pmf) of X and Y. Hint: Use a joint probability table.



all equally likely

Outcome	X	Y
1 2	1	2
2 1	2	2
1 3	1	3
3 1	3	3
2 3	2	3
3 3	3	3

Y

$f(x,y)$	2	3	$f_X(x)$
X = 1	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$
X = 2	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$
X = 3	0	$\frac{2}{6}$	$\frac{1}{3}$
$f_Y(y)$	$\frac{1}{3}$	$\frac{2}{3}$	

(b) (4 points) Find the marginal pmfs of X and Y.

$$f_X(x) = \sum_y f(x,y)$$

$$f_Y(y) = \sum_x f(x,y)$$

(c) (4 points) Find the conditional pmf of X given $Y = 3$.

$$f(x|y) = \frac{f(x,y)}{f_Y(y)}$$

For $y=3$ we get

$$f(1|3) = \frac{f(1,3)}{f_Y(3)} = \frac{1/6}{2/3} = \frac{1}{4}$$

$$f(2|3) = \frac{f(2,3)}{f_Y(3)} = \frac{1/6}{2/3} = \frac{1}{4}$$

$$f(3|3) = \frac{f(3,3)}{f_Y(3)} = \frac{2/6}{2/3} = \frac{1}{2}$$

(d) (4 points) Find $E(Y \mid X = 2)$.

$$f_{Y|X}(2 \mid 2) = \frac{f(2,2)}{f_X(2)} = \frac{1/6}{1/3} = \frac{1}{2}$$

$$f_{Y|X}(3 \mid 2) = \frac{f(2,3)}{f_X(2)} = \frac{1/6}{1/3} = \frac{1}{2}$$

$$\Rightarrow E(Y \mid X=2) = \sum_{y=2}^{3} y\, f_{Y|X}(y \mid 2)$$

$$= 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{2} = 2.5$$

(e) (5 points) Find $\mathrm{Cov}(X, Y)$.

$$E(X) = \sum_{x=1}^{3} x\, f_X(x) = 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{3} = \frac{6}{3} = 2$$

$$E(Y) = \sum_{y=2}^{3} y\, f_Y(y) = 2 \cdot \frac{1}{3} + 3 \cdot \frac{2}{3} = \frac{8}{3}$$

$$E(XY) = \sum_{Y} \sum_{X} xy\, f(x,y)$$

$$= 1 \cdot 2 \cdot \frac{1}{6} + 1 \cdot 3 \cdot \frac{1}{6} + 2 \cdot 2 \cdot \frac{1}{6} + 2 \cdot 3 \cdot \frac{1}{6} + 3 \cdot 2 \cdot 0 + 3 \cdot 3 \cdot \frac{2}{6}$$

$$= \frac{1}{6}\left(2 + 3 + 4 + 6 + 18\right) = \frac{33}{6}$$

$$\Rightarrow \mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$$

$$= \frac{33}{6} - 2 \cdot \frac{8}{3} = \frac{33-32}{6} = \frac{1}{6}$$

3. (5 points) Let X be a Normally distributed random variable with mean θ and variance $c\theta^2$. Here θ is the unknown parameter and $c(> 0)$ is a known constant. Prove that X belongs to a k-parameter exponential family and identify the value of k.

$$f(x) = \frac{1}{\sqrt{2\pi c\theta^2}} \exp\left(-\frac{(x-\theta)^2}{2c\theta^2}\right) I_{(-\infty,\infty)}(x)$$

$$= \frac{1}{\sqrt{2\pi c}} \frac{1}{\theta} \exp\left(-\frac{x^2}{2c\theta^2} + \frac{2x\theta}{2c\theta^2} - \frac{\theta^2}{2c\theta^2}\right)$$

$$= \frac{1}{\sqrt{2\pi c}} \frac{1}{\theta} \exp\left(-\frac{1}{2c}\right) \exp\left(-x^2 \frac{1}{2c\theta^2} + x\frac{1}{c\theta}\right)$$

set $h(x) = 1$ (or $h(x) = I_{(-\infty,\infty)}(x)$)

$$c(\theta) = \frac{e^{-1/2c}}{\sqrt{2\pi c}\,\theta}$$

$$t_1(x) = -x^2, \quad w_1(\theta) = \frac{1}{2c\theta^2}$$

$$t_2(x) = x, \quad w_2(\theta) = \frac{1}{c\theta}$$

$$\Rightarrow f(x) = c(\theta)\, h(x) \exp\left(t_1(x) w_1(\theta) + t_2(x) w_2(\theta)\right)$$

$\Rightarrow$ exponential family with $k = 2$.

4. Let X be a continuous random variable with pdf

$$f(x) = \begin{cases} \frac{\theta}{2}\exp(-\theta|x|) & , -\infty < x < \infty \\ 0 & , \text{ otherwise} \end{cases} \tag{1}$$

(a) (4 points) Find the expected value of X, $E(X)$

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{-\infty}^{0} x \frac{\theta}{2} \exp(\theta x)\,dx + \int_{0}^{\infty} x \frac{\theta}{2} \exp(-\theta x)\,dx$$

$$= \int_{0}^{\infty} (-x) \frac{\theta}{2} \exp(-\theta x)\,dx + \int_{0}^{\infty} x \frac{\theta}{2} \exp(-\theta x)\,dx$$

$$= -\int_{0}^{\infty} \frac{x\theta}{2} \exp(-\theta x)\,dx + \int_{0}^{\infty} \frac{x\theta}{2} \exp(-\theta x)\,dx = 0$$

(b) (4 points) Find the variance of X, $V(X)$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_{-\infty}^{0} x^2 \frac{\theta}{2} e^{\theta x}\,dx + \int_{0}^{\infty} x^2 \frac{\theta}{2} e^{-\theta x}\,dx$$

$$= \int_{0}^{\infty} (-x)^2 \frac{\theta}{2} e^{-\theta x}\,dx + \int_{0}^{\infty} x^2 \frac{\theta}{2} e^{-\theta x}\,dx$$

$$= \int_{0}^{\infty} \theta x^2 e^{-\theta x}\,dx = \frac{\Gamma(3)}{\theta^2} \int_{0}^{\infty} \underbrace{\frac{1}{\Gamma(3)} \theta^3 x^{3-1} e^{-\theta x}}_{\text{pdf of Gamma}(3, \frac{1}{\theta})}\,dx$$

$$= \frac{2!}{\theta^2} = \frac{2}{\theta^2}$$

$$\Rightarrow V(X) = E(X^2) - E(X)^2 = \frac{2}{\theta^2} - 0 = \frac{2}{\theta^2}$$

5. (6 points) Again, let X be a continuous random variable with

$$f(x) = \begin{cases} \frac{\theta}{2}\exp(-\theta|x|) & , -\infty < x < \infty \\ 0 & , \text{ otherwise} \end{cases} \tag{2}$$

Find the pdf for the random variable $Y = |X|$.

$g = |x|$ is monotone in parts

have a one-to-one function on $(-\infty, 0)$ and $(0, \infty)$

Support for Y: $[0, \infty)$

$y = g_1(x) = x$ for $0 < x < \infty$

$y = g_2(x) = -x$ for $-\infty < x < 0$

$g_1^{-1}(y) = y \Rightarrow \frac{d}{dy} g_1^{-1}(y) = 1$

$g_2^{-1}(y) = -y \Rightarrow \frac{d}{dy} g_2^{-1}(y) = -1$

$\Rightarrow f_Y(y) = \sum_{i=1}^{2} f_X(g_i^{-1}(y)) \left| \frac{d}{dy} g_i^{-1}(y) \right|$

$= \frac{\theta}{2} e^{-\theta|y|} \cdot 1 + \frac{\theta}{2} e^{-\theta|-y|} \cdot |-1|$

$= \theta e^{-\theta y}$ for $0 \le y < \infty$

$= \text{Expo}\left(\frac{1}{\theta}\right)$

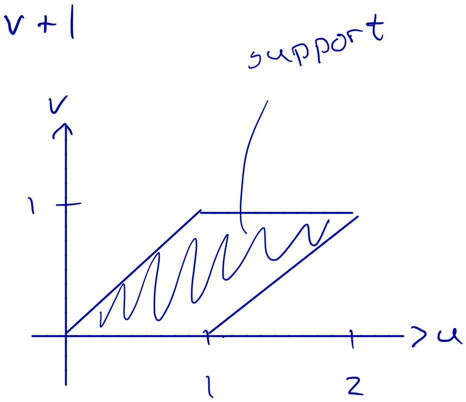
6. (8 points) Let X and Y be two independent and identically distributed Uniform$(0,1)$ random variables. Find the probability density function of $U = X + Y$

$U = X + Y$ and set $V = X$

Note that $0 < x < 1$ and $0 < y < 1$

leads to $0 < u < 2$ and $0 < v < 1$,

but also $v \le u$ and $u \le v + 1$



Inverse functions

$x = v = h_1(u,v)$

$y = u - x = u - v = h_2(u,v)$

$$J = \begin{vmatrix} \frac{\delta x}{\delta u} & \frac{\delta x}{\delta v} \\ \frac{\delta y}{\delta u} & \frac{\delta y}{\delta v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = 0 - 1 = -1$$

$$f_{X,Y}(x,y) = I_{(0,1)}(x)\, I_{(0,1)}(y)$$

$$f_{U,V}(u,v) = f_{X,Y}(h_1(u,v), h_2(u,v))\,|J|$$

$$= I_{(0,1)}(v)\, I_{(0,1)}(u-v) \cdot 1$$

$$f(u) = \int_{-\infty}^{\infty} f_{U,V}(u,v)\,dv$$

For $u < 1$: $f(u) = \int_0^u 1\,dv = v\Big|_{v=0}^{u} = u$

For $u \geq 1$: $f(u) = \int_{u-1}^1 1\,dv = v\Big|_{v=u-1}^{1} = 1 - (u-1) = 2 - u$

$$\Rightarrow f(u) = \begin{cases} u & \text{for } 0 < u < 1 \\ 2-u & \text{for } 1 < u < 2 \\ 0 & \text{otherwise} \end{cases}$$

7. (10 points) Let X and Y be independent random variables where $X \sim N(\mu_X, \sigma^2)$ and $Y \sim N(\mu_Y, \sigma^2)$. Let $U = X + Y$ and $V = X - Y$. Show that U and V are independent normal random variables and give the parameters (mean and variance) of each.

$$u = g_1(x,y) = x+y \quad \text{and} \quad v = g_2(x,y) = x-y$$

$$\Rightarrow \quad x = \frac{u+v}{2} = h_1(u,v) \quad \text{and} \quad y = \frac{u-v}{2} = h_2(u,v)$$

$$J = \begin{vmatrix} \frac{\delta x}{\delta u} & \frac{\delta x}{\delta v} \\ \frac{\delta y}{\delta u} & \frac{\delta y}{\delta v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{4} - \frac{1}{4} = -\frac{1}{2}$$

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu_X)^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(y-\mu_Y)^2}{2\sigma^2}\right)$$

$$\Rightarrow f_{u,v}(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\left(\frac{u+v}{2} - \mu_X\right)^2}{2\sigma^2}\right) \exp\left(-\frac{\left(\frac{u-v}{2} - \mu_Y\right)^2}{2\sigma^2}\right) \cdot \frac{1}{2}$$

$$= \frac{1}{2\pi \, 2\sigma^2} \exp\left(-\frac{1}{2\sigma^2}\left[\left(\frac{u+v-2\mu_X}{2}\right)^2 + \left(\frac{u-v-2\mu_Y}{2}\right)^2\right]\right)$$

$$= \frac{1}{2\pi \, 2\sigma^2} \exp\left(-\frac{1}{2\cdot 4\sigma^2}\left[\underbrace{(u+v-2\mu_X)^2 + (u-v-2\mu_Y)^2}_{=A}\right]\right)$$

$$\begin{aligned}
A = {} & u^2 + 2uv - 4u\mu_X + v^2 - 4v\mu_X + 4\mu_X^2 \\
& + u^2 - 2uv - 4u\mu_Y + v^2 + 4v\mu_Y + 4\mu_Y^2 \\
= {} & 2u^2 - 4u\mu_X - 4u\mu_Y + 2v^2 - 4v\mu_X + 4v\mu_Y + 4\mu_X^2 + 4\mu_Y^2 \\
= {} & 2\left(u^2 - 2u(\mu_X+\mu_Y) + \mu_X^2 + \mu_Y^2\right) - 2\cdot(2\mu_X\mu_Y) \\
& + 2\left(v^2 - 2v(\mu_X-\mu_Y) + \mu_X^2 + \mu_Y^2\right) + 2\cdot(2\mu_X\mu_Y) \\
= {} & 2\left((u-(\mu_X+\mu_Y)\right)^2 + 2\left(v-(\mu_X-\mu_Y)\right)^2
\end{aligned}$$

$$\Rightarrow f_{u,v}(x,y) = \frac{1}{\sqrt{2\pi}\sqrt{2}\sigma} \exp\left(-\frac{(u-(\mu_X+\mu_Y))^2}{(2\sigma)^2}\right) \frac{1}{\sqrt{2\pi}\sqrt{2}\sigma} \exp\left(-\frac{(v-(\mu_X-\mu_Y))^2}{(2\sigma)^2}\right)$$

$$= f_u(u) \cdot f_v(v) \quad \Rightarrow \text{independent and}$$

$$U \sim N(\mu_X+\mu_Y, 2\sigma^2) \quad \text{and} \quad V \sim N(\mu_X-\mu_Y, 2\sigma^2)$$

Alternative solution to 7:

Linear combinations of normally distributed random variables are also normally distributed.

Furthermore,

$$E(U) = E(X+Y) = E(X) + E(Y) = \mu_X + \mu_Y$$

$$V(U) = V(X+Y) = V(X) + V(Y) + \underbrace{2\,Cov(X,Y)}_{=0 \text{ since } X,Y \text{ are indep.}}$$

$$= 2\sigma^2$$

and

$$E(V) = E(X-Y) = E(X) - E(Y) = \mu_X - \mu_Y$$

$$V(V) = V(X-Y) = V(X) + V(Y) - 2\,Cov(X,Y)$$

$$= 2\sigma^2$$

$\Rightarrow U \sim N(\mu_X + \mu_Y, 2\sigma^2)$ and $V \sim N(\mu_X - \mu_Y, 2\sigma^2)$

U and V are independent since the are normally distributed and

$$Cov(U,V) = Cov(X+Y, X-Y)$$

$$= Cov(X,X) - Cov(X,Y) + Cov(Y,X) - Cov(Y,Y)$$

$$= V(X) - 0 + 0 - V(Y)$$

$$= \sigma^2 - \sigma^2 = 0$$

8. Consider the following hierarchical model. The conditional distribution of Y given $X = x$ is $N(x, x^2)$. The marginal distribution of X is Uniform$(0, 1)$. $\Rightarrow E(X) = \frac{1}{2}$, $V(X) = \frac{1-0}{12}$

(a) (5 points) Find $E(Y)$, i.e., the mean of the marginal distribution of Y.
Hint: You don't need to find the marginal pdf for Y

$$E(Y) = E(E(Y|X)) = E(X) = \frac{1}{2}$$

(b) (5 points) Find $V(Y)$, i.e., the variance of the marginal distribution of Y.
Hint: Again, you don't need to find the marginal pdf for Y.

$$V(Y) = V(E(Y|X)) + E(V(Y|X))$$
$$= V(X) + E(X^2) = V(X) + V(X) + (E(X))^2$$
$$= \frac{2}{12} + \left(\frac{1}{2}\right)^2 = \frac{2+3}{12} = \frac{5}{12}$$

(c) (5 points) Find $Cov(X, Y)$
Hint: You can use the fact $E(XY) = E(XE(Y \mid X))$ or equivalently $E(XY) = E(E(XY \mid X))$

$$E(XY) = E(X\,E(Y|X)) = E(X \cdot X) = E(X^2)$$
$$= V(X) + (E(X))^2 = \frac{1}{12} + \frac{1}{4} = \frac{4}{12}$$

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$
$$= \frac{4}{12} - \frac{1}{2} \cdot \frac{1}{2} = \frac{4-3}{12} = \frac{1}{12}$$

9. (8 points) Consider again the following hierarchical model. The conditional distribution of Y given $X = x$ is $N(x, x^2)$. The marginal distribution of X is Uniform$(0, 1)$.

Find the pdf of $U = \frac{Y}{X}$ and show that U and X are independent.

Hint: Use the bivariate transformation method with $U = \frac{Y}{X}$ and $V = X$.

$$u = g_1(x,y) = \frac{y}{x} \quad , \quad v = g_2(x,y) = x$$

$$\Rightarrow \quad x = v = h_1(u,v) \quad \text{and} \quad y = ux = uv = h_2(u,v)$$

$$J = \begin{vmatrix} \frac{\delta x}{\delta u} & \frac{\delta x}{\delta v} \\ \frac{\delta y}{\delta u} & \frac{\delta y}{\delta v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ v & u \end{vmatrix} = -v$$

$$f(x,y) = f(y|x)\, f(x)$$

$$= \frac{1}{\sqrt{2\pi}\, x} \exp\left(-\frac{(y-x)^2}{2x^2}\right) I_{(0,1)}(x)$$

$$\Rightarrow f(u,v) = \frac{1}{\sqrt{2\pi}\, v} \exp\left(-\frac{(uv-v)^2}{2v^2}\right) I_{(0,1)}(v) \cdot v$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2(u-1)^2}{2v^2}\right) I_{(0,1)}(v)$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(u-1)^2}{2}\right) I_{(0,1)}(v)$$

$= g(u)\, h(v) \Rightarrow U$ and V are independent, i.e. $\frac{Y}{X}$ and X are independent

$$f_U(u) = \int_0^1 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(u-1)^2}{2}\right) dv = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(u-1)^2}{2}\right)$$

$$\Rightarrow U \sim N(1,1)$$

Extra page if needed. Make sure to make clear which question(s) you are doing.