## MA395 TAKEHOME QUIZ 6 HINTS

## 1. Problem 3.11.2

In this problem we suppose that a die is rolled 6 times. Any sequence of 6 digits

$$
\left\{x_{1}, x_{2}, x_{3}, x_{4}, x_{5}, x_{6}\right\}, \quad x_{i} \in\{1,2,3,4,5,6\}
$$

could be considered a valid outcome, and there would be $6^{6}$ or 46,656 outcomes - too many to deal with without a computer. However, we can simplify the problem by considering the outcome of each roll as "success" if the result is 4 , and "failure" otherwise. Now we have a binomial model with $n=6$ and $p=1 / 6$, and there are only $2^{6}=64$ outcomes - tedious, but possible to do by hand. They are not equally likely, the probability of $k$ successes in 6 trials being

$$
P(X=k)=\binom{6}{k}\left(\frac{1}{6}\right)^{k}\left(\frac{5}{6}\right)^{6-k}, \quad k=0, \ldots, 6
$$

The first two rolls can be considered a separate experiment, with

$$
P(Y=k)=\binom{2}{k}\left(\frac{1}{6}\right)^{k}\left(\frac{5}{6}\right)^{2-k}, \quad k=0, \ldots, 2
$$

The last four rolls can also be considered a separate experiment, independent of the first two rolls, with

$$
P(Z=k)=\binom{4}{k}\left(\frac{1}{6}\right)^{k}\left(\frac{5}{6}\right)^{4-k}, \quad k=0, \ldots, 4
$$

The conditional probability can be obtained directly, if laboriously, from the definition,

$$
P(X=k \mid Y=j)=\frac{P(Y=j \text { and } Z=k-j)}{P(Y=j)}
$$

## 2. Problem 3.11.12

In this problem we are given the following joint density for $X$ and $Y$ :

$$
f_{X Y}(x, y)=2 e^{-(x+y)} \quad 0<x<y, \quad y>0
$$

and asked to find:

$$
\begin{aligned}
& P(Y<1 \mid X<1) \\
& P(Y<1 \mid X=1) \\
& f_{Y \mid x}(y) \\
& E(Y \mid x)
\end{aligned}
$$

The support is determined by the inequalities stated in the definition of $f_{X Y}(x, y)$, and is the area in the first quadrant lying above the line $y=x$ (extended to infinity).


In the above graph, the support is the area that is not shaded.
Answering the questions will involve integrating various functions over the region of support. Getting the correct answers will depend on choosing the correct limits of integration.

Recall Fubini's theorem from calculus III: If $f$ is defined on the rectangle

$$
R=\{(x, y) \mid a \leq x \leq b, c \leq y \leq d\}
$$

and is bounded on $R$ and discontinuous only on a finite number of smooth curves, then the following iterated integrals exist:

$$
\iint_{R} f(x, y) d A=\int_{a}^{b} \int_{c}^{d} f(x, y) d y d x=\int_{c}^{d} \int_{a}^{b} f(x, y) d x d y
$$

We can consider $f_{X Y}$ to be defined piecewise on R ,

$$
f_{X Y}(x, y)= \begin{cases}2 e^{-(x+y)} & x<y \\ 0 & \text { otherwise }\end{cases}
$$

With this definition, $f_{X Y}(x, y)$ is discontinuous only along the line $y=$ $x$ and the conditions for Fubini's theorem are satisfied.

Our region $R$ is defined by the rectangle $0 \leq x \leq b$ and $0 \leq y \leq d$, and the improper double integral will be the limit as $b, d \rightarrow \infty$ (which we will assume exists).

The two iterated integrals may be written as

$$
\int_{0}^{b}\left(\int_{0}^{d} f_{X Y}(x, y) d y\right) d x=\int_{0}^{d}\left(\int_{0}^{b} f_{X Y}(x, y) d x\right) d y
$$

In evaluating the inner integrals, we treat everything but the variable of integration as a constant.

We'll first consider

$$
\int_{0}^{b}\left(\int_{0}^{d} f_{X Y}(x, y) d y\right) d x
$$

When we evaluate the inner integral,

$$
\int_{0}^{d} f_{X Y}(x, y) d y
$$

we treat $x$ as a constant. If we picture the set

$$
\{(x, y) \mid x=c\}
$$

for any constant $c$, these points lie on a vertical line that crosses the $x$-axis at $x=c$.

We can consider this vertical line to be the path along which we are integrating $f_{X Y}(x, y)$.

If our support was the entire rectangle $R$, the limits of integration would be 0 to $d$,

$$
\int_{0}^{d} f_{X Y}(x, y) d y
$$

But, in our case, $f$ only has support above the line $y=x$, so our path of integration has to be modified to

$$
\{(x, y) \mid x=c \text { and } y>x\}
$$

and the integral now becomes (remember, $x$ is considered a constant for the moment):

$$
\int_{x}^{d} f_{X Y}(x, y) d y
$$

Assuming the limit exists, as $d \rightarrow \infty$ the inner integral becomes

$$
\int_{x}^{\infty} f_{X Y}(x, y) d y
$$

For the outer integral, we need to consider the set of all possible values that the constant $c$ can take in the definition of the path for the inner integral:

$$
\{(x, y) \mid x=c \text { and } y>x\}
$$

From the graph of the support, it is clear that this should be all values of $x$ in the interval $0 \leq x \leq b$, and the iterated integral will be

$$
\int_{0}^{b}\left(\int_{x}^{\infty} f_{X Y}(x, y) d y\right) d x
$$

Letting $b \rightarrow \infty$, if the improper integral exists we have finally

$$
\int_{0}^{\infty}\left(\int_{x}^{\infty} f_{X Y}(x, y) d y\right) d x=\int_{0}^{\infty}\left(\int_{x}^{\infty} 2 e^{-(x+y)} d y\right) d x
$$

You can use algebra or maple to confirm that this double improper integral evaluates to 1 , as should the integral of any joint pdf over its support.

Now consider the other possibility for the iterated integral:

$$
\int_{0}^{d}\left(\int_{0}^{b} f_{X Y}(x, y) d x\right) d y
$$

This time when we evaluate the inner integral,

$$
\int_{0}^{b} f_{X Y}(x, y) d x
$$

we treat $y$ as a constant. If we picture the set

$$
\{(x, y) \mid y=c\}
$$

for any constant $c$, these points lie on the horizontal line that crosses the $y$-axis at $y=c$.

We can consider this horizontal line to be the path along which we are integrating $f_{X Y}(x, y)$.

If our support was the entire rectangle $R$, the limits of integration would be 0 to $b$,

$$
\int_{0}^{b} f_{X Y}(x, y) d x
$$

But, in our case, $f$ only has support above the line $y=x$, so we include only the part of the horizontal line $y=c$ that lies to the left of the diagonal line $x=y$ :

$$
\{(x, y) \mid y=c \text { and } x<y\}
$$

and the integral now becomes (remember, $y$ is considered a constant for the moment):

$$
\int_{0}^{y} f_{X Y}(x, y) d x
$$

For the outer integral, we need to consider the set of all possible values that the constant $c$ can take in the definition of the path for the inner integral:

$$
\{(x, y) \mid y=c \text { and } x<y\}
$$

From the graph of the support, it is clear that this should be all values of $y$ in the interval $0 \leq y \leq d$, and the iterated integral will be

$$
\int_{0}^{d}\left(\int_{0}^{y} f_{X Y}(x, y) d x\right) d y
$$

Letting $d \rightarrow \infty$, if the improper integral exists we have finally

$$
\int_{0}^{\infty}\left(\int_{0}^{y} f_{X Y}(x, y) d x\right) d y=\int_{0}^{\infty}\left(\int_{0}^{y} 2 e^{-(x+y)} d x\right) d y
$$

Again, you can use algebra or maple to confirm that this double improper integral evaluates to 1 , as should the integral of any joint pdf over its support.

Now we have two ways of integrating the pdf over its support:

$$
\int_{0}^{\infty}\left(\int_{x}^{\infty} 2 e^{-(x+y)} d y\right) d x=\int_{0}^{\infty}\left(\int_{0}^{y} 2 e^{-(x+y)} d x\right) d y=1
$$

We can modify either of them to cover only part of the region of support. Depending on the specific part, one may be easier to work with than the other.
2.1. Part a). In this part we are asked to find

$$
P(Y<1 \mid X<1)
$$

Since the event $X<1$ has a nonzero probability, we can use the definition of the conditional probability of event $A$ given that event $B$ has occurred:

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

In this case, we will divide $P(Y<1$ and $X<1)$ by $P(X<1)$.
To evaluate $P(X<1)$, it is easier to modify

$$
\int_{0}^{\infty}\left(\int_{x}^{\infty} 2 e^{-(x+y)} d y\right) d x
$$

hint: You can obtain the marginal cdf of $X$ by modifying this integral in the following way: change all references to $x$ to references to a dummy variable $v$. Then, replace the upper limit of the outer integral with $x$. You can get the marginal pdf of $X$, which we will need later, by differentiation. The answer is

$$
f_{X}(x)=2 e^{-2 x}
$$

On the other hand, $P(Y<1$ and $X<1)$ is more easily obtained by modifying

$$
\int_{0}^{\infty}\left(\int_{0}^{y} 2 e^{-(x+y)} d x\right) d y
$$

2.2. Part b). It's possible to answer this at a glance considering the region of support.
2.3. Part c). Using the definition directly,

$$
f_{Y \mid x}(y)=\frac{f_{X Y}(x, y)}{f_{X}(x)}
$$

We obtained $f_{X}(x)$ earlier. You can verify that

$$
\int_{x}^{\infty} f_{Y \mid x}(y) d y=1
$$

(why is the lower limit $x$ instead of 0 ?) Remark: The authors write $f_{Y \mid x}(y)$ for the conditional pdf of $Y$ given $X=x$, because once the value of $X$ is fixed, we are left with a function of a single random variable, $Y$. However, you can consider a more general form $f_{Y \mid x}(x, y)$ where $x$ represents the given value of $x$, not a random variable. The answer in this case is

$$
f_{Y \mid x}=\frac{e^{-(x+y)}}{e^{-2 x}}
$$

You can verify that the integral of this function over the support of the conditional density $f_{Y \mid x}$, the interval $(x, \infty)$, is 1 for any value of $x$ in the support of the joint pdf $f_{X Y}(x, y)$. You can also use it to obtain $E(Y \mid x)$ as a function of $x$.
2.4. Part d). Once we have $f_{Y \mid x}(y)$, we can obtain $E(Y \mid x)$ directly from the definition of expected value:

$$
E(Y \mid x)=\int_{S} y \cdot f_{Y \mid x} d y
$$

where the integration is over the support of $f_{Y \mid x}$. The question at the end of the following section should serve as a hint.

## 3. Problem 3.12.19

The idea in this problem is to determine the moment-generating function for a distribution of the specified type, and use the theorems of the previous sections to find the moment-generating function of the sum of two random variables with the specified distribution (possibly, with different parameter values). Then, examine the result and see if it has the same form as the mgf's we started with. If it does, apply one of the theorems to conclude that the distribution of the sum is of the same type as the distribution of the original random variables.

