*Last Name:* *First Name:* *Student Number:*

# UNIVERSITY OF TORONTO
## Faculty of Arts and Science
## DECEMBER EXAMINATIONS 2006
## STA 247H1 F
## Duration - 3 hours

| | |
|---|---|
| 1 | |
| 2 | |
| 3 | |
| 4 | |
| 5 | |
| 6 | |
| 7 | |
| 8 | |
| 9 | |
| 10 | |
| T | |

Answer all questions in the space provided; if you run out of space, use the back of a page.

**No books or notes are allowed. No calculators are allowed.**

**You may use the facts about standard distributions and the table of normal probabilities that are on the sheets at the end of this exam.**

**For all questions, show how you obtained your answer (except when the question says no explanation is necessary).**

The ten questions are worth equal amounts.

1. You flip a nickel four times and a quarter two times. Define the following events:

| | |
|---|---|
| $A$ | The event that the nickel lands heads the first time it is flipped |
| $B$ | The event that the quarter lands heads the first time it is flipped |
| $C$ | The event that the nickel lands heads three times |
| $D$ | The event that the nickel and the quarter land heads the same number of times |

For the first four questions below, you **must** produce actual numerical answers — either decimal numbers (eg, 0.25) or simple fractions (eg, 2/8).

(a) Compute $P(C)$.

(b) Compute $P(C|A^c)$.

(c) Compute $P(D)$.

(d) Compute $P(B|D)$.

(e) Are $A$ and $B$ independent? Answer yes or no (no explanation is necessary).

(f) Are $A$ and $C$ independent? Answer yes or no (no explanation is necessary).

(g) Are $A$ and $D$ independent? Answer yes or no (no explanation is necessary).

(h) Are $A \cap B$ and $D$ independent? Answer yes or no, and explain your answer.

2. The random variables $X$ and $Y$ have the range $\{0, 1, 2\}$. The joint distribution of $X$ and $Y$ is given by the following table:

| $x$ | $y$ | $P(X = x, Y = y)$ |
|---|---|---|
| 0 | 0 | 0.2 |
| 0 | 1 | 0.1 |
| 0 | 2 | 0.1 |
| 1 | 0 | 0.0 |
| 1 | 1 | 0.0 |
| 1 | 2 | 0.2 |
| 2 | 0 | 0.3 |
| 2 | 1 | 0.1 |
| 2 | 2 | 0.0 |

Answer the following questions. You **must** produce actual numerical answers for these questions — either decimal numbers (eg, 0.25) or simple fractions (eg, 2/8).

(a) Write down tables for the marginal distributions of $X$ and of $Y$, i.e. give the values of $P(X = x)$ for all $x$, and of $P(Y = y)$ for all $y$.

(b) Write down a table for the conditional distribution of $X$ given that $Y = 2$, i.e. give the values of $P(X = x \,|\, Y = 2)$ for all $x$.

(c) Compute $E(X)$ and $E(Y)$.

(d) Compute $E(XY)$.

(e) Are $X$ and $Y$ independent? Explain why or why not.

3. Company A manufactures digital cameras with ten million pixels. Sometimes, one or more of these pixels will be non-functional. Whether one pixel is non-functional is independent of whether any other pixel is non-functional. The probability that a pixel will be non-functional is $3 \times 10^{-8}$.

   Company A tests each camera to see whether any of its pixels are non-functional. If a camera has any non-functional pixels, they discard it. However, Company B has found a way to obtain these discarded cameras, which they then sell to unsuspecting customers.

   Answer the following two questions. You may write your answers as arithmetic expressions (not actual numbers), as long as the expressions contain only numbers, not symbols (except for the standard constants $\pi$ and $e$).

   a) If you buy one of these cameras from Company B, what is the probability that more than one pixel in this camera is non-functional?

   b) If you buy one of these cameras from Company B, what is expected number of pixels in this camera that are non-functional?

4. Let $X$ and $Y$ be independent random variables. Suppose that $X$ has the geometric distribution with parameter $p_X$ and $Y$ has the geometric distribution with parameter $p_Y$. Let $Z$ be the minimum of $X$ and $Y$. Prove that $Z$ has a geometric distribution, and find the parameter, $p_Z$, of this distribution.

   Hint: This can be proved without any difficult calculations.

5. A program processes 10000 records from a databased stored on a disk. For each record, the program must read the record off the disk, which takes 12 milliseconds, and then perform some computations on the information in the record. The computer the program runs on has two processors, which operate at different speeds. If the program runs on the slow processor, processing a record takes 30 milliseconds, but if it runs on the fast processor, processing a record takes only 10 milliseconds. Every time the program resumes processing after a record has been read, it is randomly assigned to run on either the slow processor or the fast process, with equal probabilities for the each.

   Find the probability that the program will take more than 323000 milliseconds to process all 10000 records. You must find an actual numerical answer for this question — either a decimal number (eg, 0.25) or a simple fraction (eg, 2/8). An answer that is very close but not absolutely precise is acceptable. Remember that you may use the information at the back of the exam booklet. Show your work, and explain why your answer is correct.

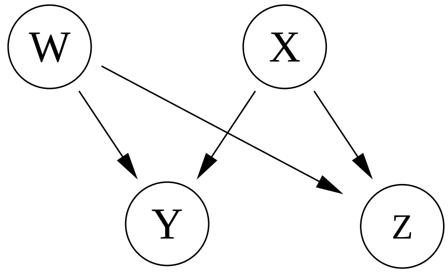
6. Suppose that the life time of a light bulb has an exponential distribution with mean 50 hours. You wish to study for 5 hours in a room light by a lamp holding such a light bulb.

   Answer the following two questions, demonstrating in detail why your answer is correct. You may write your final answers as arithmetic expressions (not actual numbers), as long as the expressions contain only numbers, not symbols (except for the standard constants $\pi$ and $e$).

   (a) Suppose you put a new light bulb in the lamp when you start studying. What is the probability that the light bulb will last at least as long as you are studying (5 hours)?

   (b) Suppose you know that a new light bulb was put in the lamp 12 hours before you start studying, and that the lamp has been on since then. What is the probability that the light bulb will last at least as long as you are studying (5 more hours)?

7. Suppose we model the joint distribution of the random variables $W$, $X$, $Y$, and $Z$ using the following causal network:




Suppose that the possible values for $W$ are 1, 2, and 3, and that the possible values for $X$, $Y$, and $Z$ are 0 and 1. Suppose that the following probabilities for these random variable conditional on the random variables that point to them in the network are known:

$$P(W=1) \;=\; P(W=2) \;=\; P(W=3) \;=\; 1/3$$

$$P(X=0) \;=\; 1/4, \qquad P(X=1) \;=\; 3/4$$

$$\begin{aligned} P(Y=1\,|\,W=1,X=0) &= 1/5, & P(Y=1\,|\,W=1,X=1) &= 1/5 \\ P(Y=1\,|\,W=2,X=0) &= 1/5, & P(Y=1\,|\,W=2,X=1) &= 1/5 \\ P(Y=1\,|\,W=3,X=0) &= 1/5, & P(Y=1\,|\,W=3,X=1) &= 7/8 \end{aligned}$$

$$\begin{aligned} P(Z=1\,|\,W=1,X=0) &= 1/2, & P(Z=1\,|\,W=1,X=1) &= 1/3 \\ P(Z=1\,|\,W=2,X=0) &= 1/3, & P(Z=1\,|\,W=2,X=1) &= 1/3 \\ P(Z=1\,|\,W=3,X=0) &= 1/3, & P(Z=1\,|\,W=3,X=1) &= 1/3 \end{aligned}$$

Write an R function called `generate` (with no arguments) that randomly generates a set of values for $W$, $X$, $Y$, and $Z$ according to the joint distribution described by this model, returning the values for the four variables as a list.

Note that minor misunderstandings of R will not be penalized for this (or the next) question.

8. Referring to the previous question: Write an R function called `prob_X_given_YZ` that takes values for $y$ and $z$ as arguments and returns an estimate of $P(X = 1 \mid Y = y, Z = z)$. This estimate should be found by calling the `generate` function from the previous question 10000 times, and using the random values generated in this way to estimate this conditional probability.

9. You flip a fair coin. If it lands heads, you roll two fair six-sided dice; if it lands tails, you roll three fair six-sided dice. Let $X$ be the sum of the numbers on all the dice that you rolled. Let $Y$ be the number of dice rolled that show the number six.

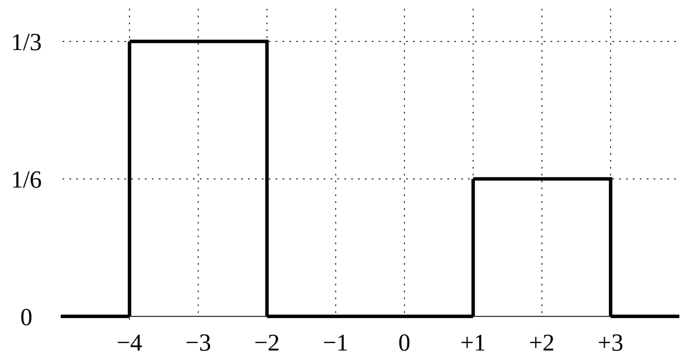
   Answer the following three questions. You **must** produce actual numerical answers for these questions — either decimal numbers (eg, 0.25) or simple fractions (eg, 2/8).

   (a) Compute $E(X)$.

   (b) Compute $E(Y)$.

   (c) Compute $VAR(Y)$.

10. Here is a graph of the probability density function for a random variable $X$:




(a) Draw a graph of the cumulative distribution function for this random variable. Label important points on the horizontal and vertical axes to make clear exactly what the function is like.

For the following three questions, you **must** produce actual numerical answers — either decimal numbers (eg, 0.25) or simple fractions (eg, 2/8).

(b) Compute $P(X \leq 0)$.

(c) Compute $E(X)$.

(d) Compute $E(X^2)$.

# Facts about standard distributions

### Binomial distribution

Parameters are $n$ and $p$. Range is the integers from 0 to $n$.

Probability mass function: $p(x) = \binom{n}{x} p^x (1-p)^{n-x}$

Mean: $E(X) = np$

Variance: $\mathrm{Var}(X) = np(1-p)$

### Geometric distribution

Parameter is $p$. Range is the integers from 1 on up.

Probability mass function: $p(x) = p(1-p)^{x-1}$

Mean: $E(X) = 1/p$

Variance: $\mathrm{Var}(X) = (1-p)/p^2$

### Poisson distribution

Parameter is $\mu$. Range is the integers from 0 on up.

Probability mass function: $p(x) = e^{-\mu}\mu^x/x!$

Mean: $E(X) = \mu$

Variance: $\mathrm{Var}(X) = \mu$

### Exponential distribution

Parameter is $\lambda$. Range is the positive real numbers.

Probability density function: $f(x) = \lambda e^{-\lambda x}$

Mean: $E(X) = 1/\lambda$

Variance: $\mathrm{Var}(X) = 1/\lambda^2$

### Normal distribution

Parameters are $\mu$ and $\sigma$. Range is the real numbers.

Probability density function: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$

Mean: $E(X) = \mu$

Variance: $\mathrm{Var}(X) = \sigma^2$