ELG3121 Final Exam

Total 60 points; 3 hours; closed-book; calculators allowed

July 18, 2005

Questions are not ordered according to their level of difficulty.

Question 1 (7 points) *The pdf of random variable Y is given in Figure 1.*

1. (3 points) Find the mean of Y.
2. (4 points) Find the variance of Y.

Question 2 (7 points) *The pdf of random variable Y is given in Figure 1. Random variable X is defined as*

$$X := \begin{cases} 1, & \text{if } Y < 3.5 \\ Y^2, & \text{if } Y \geq 3.5 \end{cases}$$

1. (4 points) Determine the cdf of X.
2. (3 points) Determine the pdf of X.

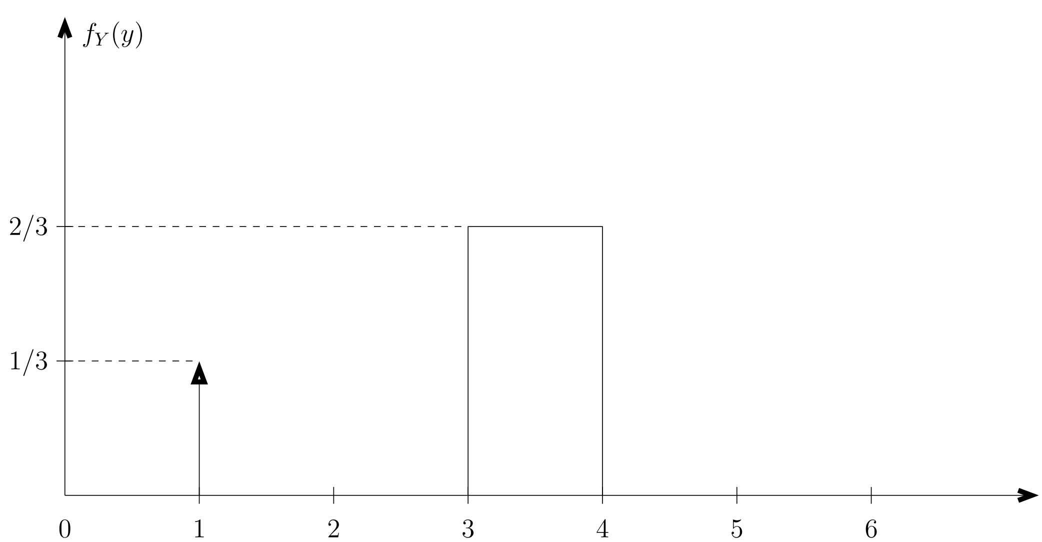

Figure 1: A probability density function

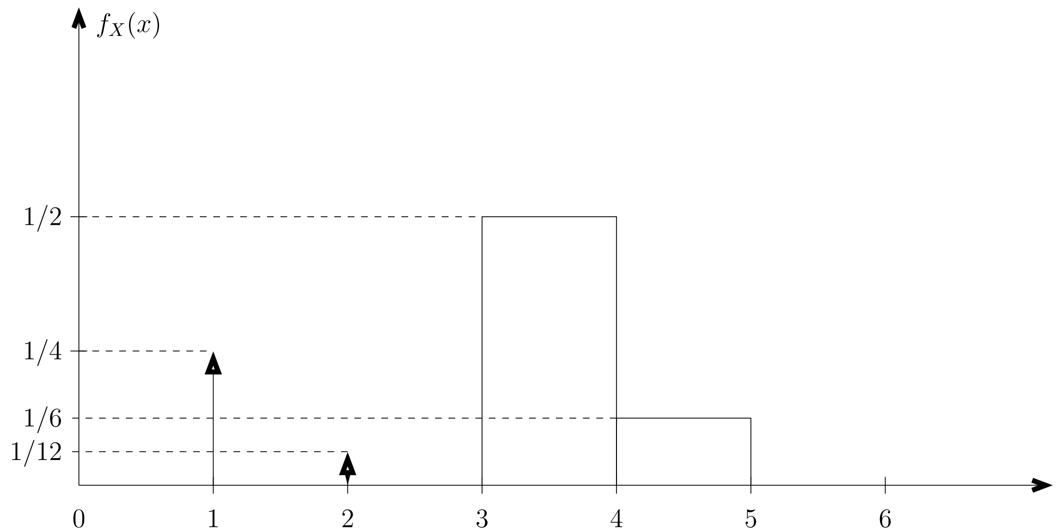

Figure 2: A probability density function

Question 3 (8 points) *Suppose that you are only allowed to perform the following two random experiments.*

- *Experiment 1, which generates a random number following distribution given in Figure 1.*
- *Experiment 2, which is tossing a fair coin.*

You are allowed to perform the two random experiments as many times as you want.

1. (6 points) By performing the two random experiments and any arithmetic computation, generate a random variable with distribution given in Figure 2. Describe your procedure (flow-chart or pseudo-code is recommended).
2. (2 points) Suppose that now you are not allowed to do Experiment 2 but still allowed to do Experiment 1 and any arithmetic computation. Can you still generate a random variable with distribution in Figure 2? Sketch your procedure if you can, or explain why this is impossible.

Question 4 (8 points) *It is hypothesized that a coin is biased, where HEAD is suspected to occur with probability* 0.6 *when the coin is tossed. You are going to toss the coin* 100 *times to test this hypothesis.*

Suppose that you set the significance level to be 1%, then what criterion should you use, in terms of the number of HEADs observed, to reject the hypothesis?

Question 5 (7 points) *The outcome of a random experiment is a Gaussian random variable X with unknown mean. The variance of X is however known to be* 25. *Now you are going to estimate the mean μ of random variable X by doing the experiments n times. Denote by $X_1, X_2, \ldots, X_n$ the outcome of the n random experiments. Denote by M_n the average of $X_1, X_2, \ldots, X_n$. You will use M_n as the estimate of μ.*

What is the minimal number n of of experiments that you need to perform so that the probability of the true mean falling into interval $(M_n - 0.1, M_n + 0.1)$ is no smaller than 0.98?

Question 6 (8 points) *Verify whether the following statements are correct. Disprove or construct counter-examples to show a statement is false; or give sufficient justification to show a statement is correct.*

1. (2 points) If the correlation coefficient between random variables X and Y is 0, then X and Y are independent.
2. (2 points) For any vector random variable (X, Y), $VAR[X]+VAR[Y] = VAR[X+Y]$.
3. (2 points) Independent random variables X and Y are both Gaussian, then there exists no region $\mathcal{C}$ of area 1 on the x-y plane such that conditioned on the event $(X, Y) \in \mathcal{C}$, X and Y are independent.
4. (2 points) Suppose that X is a Gaussian random variable, A is an arbitrary random variable independent of X, and Y is defined by $Y = AX$. Then Y is Gaussian.

Question 7 (8 points) *Suppose that the probability that a senior woman has Cancer X is* 0.1*. Now there is new test for Cancer X. If a senior woman has Cancer X, the test will show "positive" with probability* 0.9*; and if the woman does not have Cancer X, the test will show "positive" with probability* 0.3*.*

The test result for Jennifer, a senior woman, is positive. What is the probability that she has Cancer X?

Question 8 (7 points) *Over any time interval of duration* 1 *hour, the number of emails arriving at a computer is modelled as a Poisson random variable with mean* 5*. Suppose that during the past time interval* $[0, 2]$ *(in unit of hour), 5 emails have arrived.*

1. (3 points) Determine the probability that the total number of emails arriving during the interval $[0, 5]$ equals 20.
2. (4 points) Determine the probability that the 6^{th} email arrives at time no earlier than 5.

Appendix

k	$x = Q^{-1}(10^{-k})$
1	1.2815
2	2.3263
3	3.0902
4	3.7190
5	4.2649
6	4.7535

Table 1: Table of the inverse of Q-function (evaluated at 10^{-k}).

Degree of freedom	Threshold for $\alpha = 5\%$	Threshold for $\alpha = 1\%$
1	3.84	6.63
2	5.99	9.21
3	7.81	11.35
4	9.49	13.28
5	11.07	15.09
6	12.59	16.81

Table 2: Table of the threshold values of the Chi-Square test with various degrees of freedom and with significance level $\alpha = 5\%, 1\%$.