Probability Review Test

August 15, 2017

Statistical Learning and Graphical Models

University of Virginia, Fall 2017

The recommended time is 90 mins, but you can take longer if you need to.

Questions marked by a * are required for graduate students but bonus for undergraduates.

Due date: Tuesday, 8/22, 5:00 pm

Problem 1

(18 points) True or False? A statement is True if and only if it is always true. Each correct choice counts +3 points, whereas an incorrect choice counts -1 point.

a) True ☐ False ☐ For two events A and B, $P(A \cup B) = P(A) + P(B)$.

b) True ☐ False ☐ For two events A and B, $P(A \cap B) = P(A)P(B)$.

c) True ☐ False ☐ For random variables X and Y, we have $E\left[X + Y^2\right] = E[X] + (E[Y])^2$.

d) True ☐ False ☐ The following is a valid probability mass function (pmf) for a random variable X.

$$p_X(i) = (e-1)\, e^{-i} \qquad \text{for } i = 1, 2, \cdots$$

e) True ☐ False ☐ For random variables X and Y, we have $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$.

f) True ☐ False ☐ Let A and B be events with probabilities $P(A) = 3/4$ and $P(B) = 2/3$. It is possible to have $P(A \cap B) = 1/6$.

Problem 2

(2×5 pts)
Let X be a random variable defined as

$$X = \begin{cases} 0, & \text{with probability } 1/2 \\ 1, & \text{with probability } 1/4 \\ 2 & \text{with probability } 1/4 \end{cases}$$

Find

a) the expected value of X, and

b) the variance of X.

Problem 3

(6 + 4 pts) Suppose X is a uniform random variable over the interval $\left[0, \frac{3}{2}\right]$. Let Y be defined as $Y = \lfloor X \rfloor$, where $\lfloor x \rfloor$ is defined as the largest integer that is less than or equal to x.

a) Find the pmf of Y.

b) What is $E\left[Y|X = 1.25\right]$?

Problem 4

(15 pts) Suppose $X_1, X_2, \cdots, X_n$ are independent discrete random variables. Assume that each X_i has mean μ_i and variance σ_i^2. Let

$$X = \sum_{i=1}^{n} X_i$$

and show that

$$\boxed{\mathrm{Var}\left[X\right] = \sum_{i=1}^{n} \mathrm{Var}\left[X_i\right] = \sum_{i=1}^{n} \sigma_i^2.}$$

If, in addition, all X_i have the same distribution, $X_1, X_2, \cdots, X_n$ is called a sequence of *independent, identically distributed (iid)* random variables. Assuming $X_1, X_2, \cdots, X_n$ are iid and each X_i has mean μ and variance σ^2, what is $\mathrm{Var}\left[X\right]$?

Problem 5

(3+4+5 pts)

There are two dice in a bag. One has one red face, and the other two red faces. One of the dice is drawn at random from the bag, each die having an equal chance of being drawn. The selected die is repeatedly rolled.

a) What is the probability that red shows on the first roll?

b) Given that red shows on the first roll, what is the conditional probability that red shows on the second roll?

c) Given that red shows on the first two rolls, what is the conditional probability that the selected die has red on two faces?

Problem 6

(8+7 pts) Suppose X and Y are two random variables with a joint probability mass function (pmf). Define

$$g(y) = E[X|Y=y]$$
$$E[X|Y] = g(Y)$$

which implies that

$$E[E[X|Y]] = E[g[Y]].$$

a) Prove that

$$E[E[X|Y]] = E[X].$$

b) Verify that the equality holds for X and Y with the following pmf:

$p_{X,Y}(x,y)$	$x=0$	$x=1$
$y=0$	1/4	1/4
$y=1$	1/2	0

Problem 7

(10 pts) Consider random variables $X \sim N(1,1)$ and $Y \sim N\left(0,3^2\right)$. Suppose X and Y are jointly Gaussian and that the correlation coefficient of X and Y is ρ. Determine $P(2X+3Y-2 \geq 2)$. (Your answers should depend on ρ.)

Problem 8

(10 pts, *) Let $\Theta \sim \text{Uni}[-\pi/2, \pi/2]$. Find the pdf of Y if $Y = \tan\Theta$.