University of Illinois Fall 2018

ECE 313: Final Exam

Monday, December 17, 2018
1:30 p.m. — 4:30 p.m.

1. **[8+6+8 points]** Suppose two fair dice are rolled. Consider the following events:

   $A =$ "Second die shows a strictly larger number than the first die"

   $B =$ "Sum of the dice equals 6"

   $C =$ "Second die shows a number that is twice the number on the first die"

   (a) Find $P(A)$.
   Solution: The set A contains half the remaining outcomes after subtracting the 6 doubles. Thus
   $$P(A) = \frac{36-6}{2\times 36} = \frac{15}{36}$$

   (b) Find $P(B|A)$.
   Solution: Only two outcomes $(1,5)$ and $(2,4)$ contribute to AB. Therefore
   $$P(AB) = \frac{2}{36} \implies P(B|A) = \frac{P(AB)}{P(A)} = \frac{2}{15}.$$

   (c) Are events A and C independent? Explain.
   Solution: The event C is a subset of the event A. Therefore
   $$P(AC) = P(C) \neq P(A)P(C)$$
   which means that A and C are not independent.

2. **[8+6+8 points]** The three parts are unrelated.

   (a) Suppose X is a binomial random variable with parameters $n = 16$ and $p = 1/2$. Using the Central Limit Theorem, express $P(X \geq 10)$ in terms of the Q function **without** using the continuity correction.
   Solution: We note that $E[X] = np = 8$ and $\mathrm{Var}(X) = np(1-p) = 16(1/2)(1/2) = 4$. Using the CLT, we approximate X by $\tilde{X} \sim \mathcal{N}(E[X], \mathrm{Var}(X))$. Therefore, we have:
   $$P(X \geq 10) \approx P(\tilde{X} \geq 10) = P\left(\frac{\tilde{X}-8}{\sqrt{4}} \geq \frac{10-8}{\sqrt{4}}\right) = Q(1).$$

   (b) Assume that people show up from the corner of a near building to your place according to a Poisson process with rate $\lambda = 2$ people per hour. Find the probability of at least 3 people showing up in the next 2 hours. You can leave your answer in terms of e, the base of natural logarithm, e.g. $2e^{-1}$.
   Solution:
   $$\begin{aligned} P(N(2) \geq 3) &= 1 - P(N(2)=0) - P(N(2)=1) - P(N(2)=2) \\ &= 1 - \sum_{k=0}^{2} e^{-4}\frac{4^k}{k!} = 1 - 13e^{-4}. \end{aligned}$$

(c) Suppose that in your kitchen there is a box with n apples. You particularly like apples, therefore every day you remove an apple from the box and you eat it. To avoid a fruit shortage in your home, your mother replaces every day the fruit that you ate by an apple with probability p or by an orange with probability $1-p$. Find the expected number of days till there are no more apples in the box.

Solution: Each day, an apple is totally removed from the box with probability $1-p$ and the number of apples decreases by 1. Also, if at a particular day the box contains k apples, the box will contain at most k apples in any subsequent day, since you definitely eat an apple every day. The number of days required to finish the apples in the box is a negative binomial random variable with parameters n and $1-p$. Therefore, the expected number of days to eat all apples is $n/(1-p)$.

3. **[10+4 points]** Two sensors are used to detect whether a patient has sepsis. The first sensor outputs a value X and the second sensor outputs a value Y. Both outputs have possible values $0, 1, 2$, with larger numbers tending to indicate that the patient has sepsis. Suppose

	$X=0$	$X=1$	$X=2$
H_1	0.1	0.3	0.6
H_0	0.6	0.2	0.2

	$Y=0$	$Y=1$	$Y=2$
H_1	0.1	0.1	0.8
H_0	0.7	0.2	0.1

given one of the hypotheses is true, the sensors provide conditionally independent readings, i.e., $P(X,Y|H_i) = P(X|H_i)P(Y|H_i)$ for $i = 0, 1$.

(a) Find the likelihood matrix for the observation (X,Y) and describe the ML decision rule for this problem.

Solution: The likelihood matrix for observation (X,Y) is the following The ML de-

(X,Y)	$(0,0)$	$(0,1)$	$(0,2)$	$(1,0)$	$(1,1)$	$(1,2)$	$(2,0)$	$(2,1)$	$(2,2)$
H_1	0.01	0.01	$\underline{0.08}$	0.03	0.03	$\underline{0.24}$	0.06	$\underline{0.06}$	$\underline{0.48}$
H_0	$\underline{0.42}$	$\underline{0.12}$	0.06	$\underline{0.14}$	$\underline{0.04}$	0.02	$\underline{0.14}$	0.04	0.02

cisions are indicated by the underlined elements. The larger number in each column is underlined. Note that the row sums are both 1.

(b) Find $p_{\text{false alarm}}$ for the ML rule found in part (a).

Solution: For the ML rule, $p_{\text{false alarm}}$ is the sum of the entries in the row for H_0 in the likelihood matrix that are not underlined. So $p_{\text{false alarm}} = 0.06 + 0.02 + 0.04 + 0.02 = 0.14$.

4. **[8+6+6 points]** Let X and Y be independent random variables, both with mean 0 and variance 1. Define the random variables

$$V = 2X + 3Y \quad \text{and} \quad W = X - Y.$$

(a) Compute the linear MMSE estimator $\hat{E}[V|W]$.

Solution:

$$\hat{E}[V|W] = E[V] + \frac{\text{Cov}(V,W)}{\text{Var}(W)}(W - E[W]) = \frac{\text{Cov}(V,W)}{\text{Var}(W)}W.$$

We now compute $\text{Cov}(V,W) = E[(2X+3Y)(X-Y)] = 2\text{Var}(X) - 3\text{Var}(Y) = -1$ and $\text{Var}(W) = \text{Var}(X-Y) = \text{Var}(X) + \text{Var}(Y) = 2$. Therefore,

$$\hat{E}[V|W] = -\frac{1}{2}W.$$

(b) Compute the Mean Square Error $E\left[(V - \hat{E}[V|W])^2\right]$.

Solution:

$$E\left[(V - \hat{E}[V|W])^2\right] = \text{Var}(V)(1-\rho_{V,W}^2) = 13\left(1 - \frac{(-1)^2}{13\cdot 2}\right)$$
$$= 13 - \frac{1}{2} = \frac{25}{2} = 12.5.$$

(c) Assume instead that W is defined as $W = X - aY$ for some real a. Can V and W be uncorrelated for some value of a? Justify your answer.
Solution: Setting $\text{Cov}(V,W) = 0$, we obtain:

$$0 = \text{Cov}(V,W) = E[(2X+3Y)(X-aY)] = 2E[X^2] - 3aE[Y^2] = 2 - 3a.$$

Therefore, V, W are uncorrelated for $a = 2/3$.

5. **[4+8+6 points]** Suppose that X and Y have a joint density function f given by

$$f_{X,Y}(u,v) = \begin{cases} 1/\pi, & u^2+v^2 < 1, \\ 0, & u^2+v^2 \geq 1. \end{cases}$$

(a) Are X and Y independent?
Solution: X and Y are dependent, because the support is not a product set using swap test. Take $(0,1)$ and $(1,0)$, both points are within the support. However, after swap $(1,1)$ is not within the support.

(b) Compute the probability density f_X for X.
Solution: When $|u| \leq 1$,

$$f_X(u) = \int_{-\sqrt{1-u^2}}^{\sqrt{1-u^2}} \frac{1}{\pi} dv = \frac{2\sqrt{1-u^2}}{\pi},$$

for $|u| > 1$, $f_X(u) = 0$. The support for $f_X(u)$ is $(-1,1)$.

(c) What is $P(|Y| + |X| < 1)$?
Solution: The area of the region for $|Y| + |X| < 1$ is 2. Therefore,

$$P(|Y| + |X| < 1) = \frac{2}{\pi}$$

6. **[8 points]** Suppose X_1, X_2, ... X_n is a sequence of random variables such that each X_k has finite mean μ and variance 2, and $\text{Cov}(X_i, X_j) = -\frac{1}{n}$ for $i \neq j$. Let $S_n = \sum_{k=1}^n X_k$. For a given $\delta > 0$, use Chebychev inequality to obtain an upper bound of

$$P\left\{\left|\frac{S_n}{n} - \mu\right| \geq \delta\right\}.$$

Solution: The mean of $\frac{S_n}{n}$ is given by

$$E\left[\frac{S_n}{n}\right] = E\left[\frac{\sum_{k=1}^n X_k}{n}\right] = \frac{\sum_{k=1}^n E[X_k]}{n} = \frac{n\mu}{n} = \mu.$$

The variance of $\frac{S_n}{n}$ is given by:

$$\begin{aligned}\mathrm{Var}\left(\frac{S_n}{n}\right) &= \mathrm{Var}\left(\frac{\sum_{k=1}^n X_k}{n}\right) = \frac{\mathrm{Cov}\left(\sum_{k=1}^n X_k, \sum_{k=1}^n X_k\right)}{n^2} \\ &= \frac{\sum_{k=1}^n \mathrm{Var}(X_k) + \sum_{i\neq j}\mathrm{Cov}(X_i, X_j)}{n^2} = \frac{2n + n(n-1)(-\frac{1}{n})}{n^2} = \frac{n+1}{n^2}\end{aligned}$$

Using Chebyshev,

$$P\left\{\left|\frac{S_n}{n} - \mu\right| \geq \delta\right\} \leq \frac{\mathrm{Var}\left(\frac{S_n}{n}\right)}{\delta^2} = \frac{n+1}{n^2\delta^2}.$$

7. **[8+6 points]** Consider a 6×6 square board, which consists of 36 squares in 6 rows and 6 columns.

   (a) How many different rectangles, comprised entirely of the board squares, can be drawn on the board? *Hint:* there are 7 horizontal and 7 vertical lines on the board.
   Solution: A rectangle is uniquely described by the pair of horizontal lines and the pair of vertical lines that form its sides. Since there are $\binom{7}{2} = \frac{7\times 6}{2} = 21$ choices for the pair of horizontal lines, an, similarly, 21 choices for the pair of vertical lines, there are $21 \times 21 = 441$ rectangles

   (b) One of the rectangles you counted in part (a) is chosen at random. What is the probability that it is a square?
   Solution: The number of square shaped rectangles is $(7-k)^2$, Hence, the number of square shaped rectangles is $1^2 + 2^2 + 3^2 + \cdots + 6^2 = 7 \times 13$. So the probability of getting a square shaped rectangle is $\frac{13}{63}$.

8. **[10 points]** Given independent random variables X, Y and B. The random variable X has a uniform distribution over the interval $[0, 20]$, Y has a uniform distribution over the interval $[0, 10]$, and B has a Bernoulli distribution with $p = \frac{2}{3}$. Let $Z = BX + (1-B)Y$. Find $P(B = 1|Z > 5)$.

   Solution: Using Bayes' formula,

$$\begin{aligned}P(B=1|Z>5) &= \frac{P(B=1, Z>5)}{P(Z>5)} = \frac{P(B=1,Z>5)}{P(B=1,Z>5)+P(B=0,Z>5)} \\ &= \frac{P(B=1,X>5)}{P(B=1,X>5)+P(B=0,Y>5)} \\ &= \frac{P(B=1)P(X>5)}{P(B=1)P(X>5)+P(B=0)P(Y>5)} \\ &= \frac{\frac{2}{3}\cdot\frac{3}{4}}{\frac{2}{3}\cdot\frac{3}{4}+\frac{1}{3}\cdot\frac{1}{2}} = \frac{3}{4}.\end{aligned}$$

9. **[6+8+6 points]** Let X and Y be jointly Gaussian random variables with $\mu_X = 0$, $\mu_Y = 1$, $\sigma_X^2 = 4$, $\sigma_Y^2 = 1$.

   (a) If $\rho = \frac{1}{8}$, find $P(X + 2Y > 2)$.
   Solution: Since $X + 2Y$ is a linear combination of jointly Gaussian random variables, it is a Gaussian random variable. $E(X + 2Y) = \mu_X + 2\mu_Y = 2$. Since a Gaussian random variable is symmetric with respect to its mean, $P(X + 2Y > 2) = P(X + 2Y > E(X + 2Y)) = 0.5$.

(b) If $\rho = \frac{1}{2}$, find $E[Y|X]$.
Solution: Since X and Y are jointly Gaussian random variables,

$$E[Y|X] = \hat{E}[Y|X] = \mu_Y + \frac{\rho\sigma_Y}{\sigma_X}(X - \mu_X) = 1 + \frac{X}{4}.$$

(c) If $\rho = 0$, find $f_{Y|X}(v|u)$.
Solution: Since X and Y are jointly Gaussian random variables and $\rho = 0$, X and Y are independent. Hence

$$f_{Y|X}(v|u) = f_Y(v) = \frac{1}{\sqrt{2\pi}}e^{-\frac{(v-1)^2}{2}},$$

since Y is a Gaussian random variable with mean μ_Y and variance σ_Y^2.

10. **[8+6+8 points]** Let X be an exponentially distributed random variable with parameter 1. Let $Y = \lfloor \frac{X}{2} \rfloor$, which is the integer part of $\frac{X}{2}$.

(a) Find the distribution of Y.
Solution: Since Y is a discrete-type random variable with support on the non-negative integers, the pmf of Y is

$$p_Y(k) = P(k \le \frac{X}{2} < k+1) = P(2k \le X < 2k+2) = \int_{2k}^{2k+2} e^{-u}du = e^{-2k}(1 - e^{-2}),$$

for integers $k \ge 0$, and $p_Y(k) = 0$ for other k.

(b) Find a function g such that, if U is uniformly distributed over the interval $[0, 1]$, $g(U)$ has the distribution of X.
Solution: Since $F_X(c) = 1 - e^{-c}$ for $c \ge 0$ and $F(c) = 0$ for $c < 0$. We'll let $g(u) = F^{-1}(u)$. Since F is strictly and continuously increasing over the support, if $0 < u < 1$ then the value c of $F^{-1}(u)$ is such that $F(c) = u$. That is, we would like $1 - e^{-c} = u$ which is equivalent to $e^{-c} = 1 - u$, or $c = -\ln(1 - u)$. Thus, $F^{-1}(u) = -\ln(1 - u)$. So $g(u) = -\ln(1 - u)$ for $0 < u < 1$.

(c) Let Z be another exponentially distributed random variable with parameter 1. The random variables X and Z are independent. Let $T = \min(X, Z)$. Find the failure rate function of T.
Solution: By the independence of X and Z,

$$P(T > t) = P(X > t \text{ and } Z > t) = P(X > t)P(X > t) = e^{-t}e^{-t} = e^{-2t},$$

which is an exponential random variable with $\lambda = 2$. Hence the failure rate

$$h_T(t) = \frac{f_T(t)}{1 - F_T(t)} = \frac{2e^{-2t}}{e^{-2t}} = 2.$$

11. **[30 points]** (3 points per answer)
In order to discourage guessing, 3 points will be deducted for each incorrect answer (no penalty or gain for blank answers). A net negative score will reduce your total exam score.

(a) Consider the events such that $P(ABC) = P(B)P(AC) > 0$ and $P(BC) = P(B)P(C)$.

TRUE	FALSE	
☐	☐	$P(A\|BC) = P(A\|C)$.
☐	☐	$P(B\|AC) > P(B\|C)$.
☐	☐	If $P(A) < P(C)$, then $P(A\|C) > P(C\|A)$.

Solution: True, False, False

(b) Suppose a coin shows head with unknown probability p. Three experiments are conducted. In the first experiment, the coin is flipped 10 times and the number of heads is denoted by X. In the second experiment, the coin is flipped another 10 times and the number of heads is denoted by Y. In the third experiment, the coin is flipped another 20 times and the number of heads is denoted by Z.

TRUE	FALSE	
☐	☐	Given $X = 2$, the ML estimate of p is 0.2.
☐	☐	Given $X = 2$ and $Y = 4$, the ML estimate of p is $\frac{0.2+0.4}{2} = 0.3$.
☐	☐	Given $X = 2$ and $Z = 5$, the ML estimate of p is $\frac{0.2+0.25}{2} = 0.225$.

Solution: True,True, False,

(c) The following parts are independent.

TRUE	FALSE	
☐	☐	Suppose $X \sim \text{Geo}(p)$, $P(X > k) = (1-p)^k, k \geq 1$. Then $P(X \geq k) = (1-p)^{k+1}$.
☐	☐	Given X and Y are random variables, we always have $E[(Y - E[Y\|X])^2] < E[(Y - \hat{E}[Y\|X])^2]$.

Solution: False, False

(d) Suppose U_1, U_2, ... U_n is a sequence of i.i.d. random variables such that each U_k has a uniform distribution over $[0, c]$. Consider the product $\prod_{k=1}^{n} U_k$ as $n \to \infty$.

TRUE	FALSE	
☐	☐	If $c = 2$, $P(\prod_{k=1}^{n} U_k > \delta) \to 0$ as $n \to \infty$ for any $\delta > 0$.
☐	☐	If $c = 3$, $P(\prod_{k=1}^{n} U_k > \delta) \to 0$ as $n \to \infty$ for any $\delta > 0$.

Solution: True, False