University of Illinois Fall 2015

ECE 313: Final Exam

Tuesday, December 15, 2015
7 p.m. — 10 p.m.
Room ECEB 1015

1. [**12 points**] The three parts of this problem are unrelated.

   (a) Count the number of distinct 7-letter sequences that can be obtained from permutating the letters in the word "darling".

   Solution: The string "darling" contains 7 characters. Therefore, the number of ways in which one can arrange the letters is 7!.

   (b) Count the number of distinct 6-letter sequences that can be obtained from permutating the letters in the word "banana".

   Solution: The string "banana" contains 3 a, 2 n and 1 b. Therefore, the number of distinct words is $\frac{6!}{3!2!1!} = 60$.

   (c) There are 6 socks in a drawer: 2 black, 2 brown, and 2 gray. Suppose there are 3 siblings present, and one at a time, they each randomly grab two socks out of the drawer, without replacement.
   i) Find the probability that each sibling draws a matching pair of socks.
   ii) Find the probability that no sibling draws a matching pair of socks.
   Solution:
   i) The first sibling draws a matching pair of socks with probability 1/5 (the first sock can be anything while the second has to match the first one). Given that the first sibling drew a matching pair of socks, the second sibling draws a matching pair of socks with probability 1/3 (again, the first sock can be anything while the second has to match the first one). Given that the first and second siblings drew matching pairs of socks, the third sibling draws a matching pair of socks with probability 1. Therefore, the probability that each sibling draws a matching pair of socks is $(1/5) \times (1/3) = 1/15$.
   ii) The first sibling draws non-matching socks with probability 4/5 (the first sock can be anything (say black) while the second one can be anything but black). At this point, we are left with 2 socks of the same color and 2 other socks of different colors. For the sake of simplicity, assume that we are left with 2 brown (R), 1 black (B), and 1 gray (G). Now, the second sibling can draw $\{R, G\}$, $\{G, R\}$, $\{R, B\}$, or $\{B, R\}$ each with probability 1/6. Notice that $\{G, B\}$ and $\{B, G\}$ are not allowed because the second sibling has to leave a non-matching pair of socks for the third sibling. Thus, the second sibling draws non-matching socks with probability $4/6 = 2/3$. Given this, the third sibling draws non-matching socks with probability 1. Therefore, the probability that no sibling draws a matching pair of socks is $(4/5) \times (2/3) = 8/15$.

2. [**18 points**] Alice transmits a single bit X to Bob over a noisy channel. For each transmission, the output is equal to X with probability p and $1 - X$ with probability $1 - p$, independently of other transmission attempts. Alice attempts n transmissions

and Bob, on the other end of the channel, observes $Y_1, \ldots, Y_n$, the n outputs of the channel, and computes $S_n = Y_1 + \cdots + Y_n$.

(a) Assuming that $X = 1$, find the pmf of S_n.
Solution: When $X = 1$, Y_i is a Bernoulli random variable with probability p. Therefore, S_n is a Binomial random variable with parameters n and p and its pmf is given by

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

(b) Assuming that $X = 1$, $E[S_n] = 12$ and $Var(S_n) = 3$, find the value of p and n.
Solution: From part (a), $E[S_n|X = 1] = np = 4$ and $\mathrm{Var}(S_n|X = 1) = np(1-p) = 3$. Solving for n and p, we get that $n = 16$ and $p = 3/4$.

(c) Assuming that $X = 1$. Find the maximum likelihood estimate of n given that $S_n = 10$ is observed.
Solution: From part (a), the likelihood of $S_n = 10$ is

$$f(n) = P(S_n = 10) = \binom{n}{10} p^k (1-p)^{n-k} = \binom{n}{10} (1-p)^n \left(\frac{p}{1-p}\right)^k,$$

for $n \geq 10$ and equal to 0 for $n < 10$. Observe that

$$\binom{n}{10} = \frac{n(n-1)\cdots(n-9)}{10!}.$$

increases with n, whereas $(1-p)^n$ decreases exponentially with n. Thus, when n is small, $\binom{n}{10}$ dominates $(1-p)^n$ and therefore $f(n)$ increases with n until it reaches a maximum $f(n^*)$ at n^*, the maximum likelihood estimate of n. To find n^*, we have to the smallest n^* such that $f(n^*+1) - f(n^*) \leq 0$. This happens when $n^* = \lfloor \frac{10}{p} \rfloor$. Observe that the maximum likelihood estimate in this case need not be unique. For instance, for $p = 1/2$, since $f(19) = f(20)$, both $n = 19$ and $n = 20$ have the same likelihood (i.e., either one of them can be used as the maximum likelihood estimate of n).

(d) Assuming that X is a Bernoulli random variable with parameter $1/2$, find the pmf of S_n.
Solution: Given that $X = 0$, each Y_i is a Bernoulli random variable with parameter $1-p$ and thus, S_n is a binomial random variable with parameters n and $1-p$. Given that $X = 1$, each Y_i is a Bernoulli random variable with parameter p and thus, S_n is a binomial random variable with parameters n and p. Using the law of total probability, the pmf of S_n is given by

$$\begin{aligned}
P(S_n = k) &= P(S_n = k|X = 1)P(X = 1) + P(S_n = k|X = 0)P(X = 0) \\
&= \frac{1}{2}\binom{n}{k} p^k (1-p)^{n-k} + \frac{1}{2}\binom{n}{k} p^{n-k} (1-p)^k \\
&= \frac{1}{2}\binom{n}{k} \left(p^k (1-p)^{n-k} + p^{n-k} (1-p)^k\right).
\end{aligned}$$

3. **[22 points]** Consider the experiment of rolling two fair dice. Define the two events $A = \{\text{both dice show a different number}\}$ and $B = \{\text{sum of numbers showing is} \leq 4\}$.

   (a) Obtain $P(B|A)$ and determine if A and B are independent (explain).

   Solution: Outcomes are equally likely, so $P(\text{event}) = \frac{|\text{event}|}{|\Omega|}$. Therefore, by the definition of conditional probability, $P(B|A) = \frac{P(AB)}{P(A)} = \frac{|AB|}{|A|} = \frac{4}{30} = \frac{2}{15}$. $P(B) = \frac{|B|}{|\Omega|} = \frac{6}{36} \neq P(B|A)$, hence they are not independent.

   (b) Repeat the experiment 10 times and let X be the number of times that event A occurs. Obtain the pmf of X.

   Solution: In each experiment trial, event A either occurs, or it doesn't occur, hence $X \sim Binomial(10, 5/6)$, because $P(A) = 5/6$. Therefore, $p_X(k) = \binom{10}{k}\left(\frac{5}{6}\right)^k\left(\frac{1}{6}\right)^{10-k}$ for $k \in \{0, 1, \ldots, 10\}$.

   (c) Let $Y = -3X$, with X defined as above. Obtain $E[Y]$, $Var(Y)$ and $E[Y^2]$.

   Solution: By linearity of expectation $E[Y] = E[-3X] = -3E[X] = -3(10)\frac{5}{6} = -25$.
   $Var(Y) = Var(-3X) = 9Var(X) = 9\left(10\frac{5}{6}\frac{1}{6}\right) = \frac{25}{2}$.
   $E[Y^2] = Var(Y) + \mu_Y^2 = \frac{25}{2} + 25^2 = 637.5$.

   (d) Are X and Y, as defined above, independent? Explain.

   Solution: If indepedent, then $P\{X = j, Y = k\} = P\{X = j\}P\{Y = k\}$ for all j, k. However, $P\{X = 1, Y = -6\} = P\{X = 1, X = 2\} = 0 \neq P\{X = 1\}P\{Y = -6\}$. So they are dependent.

4. **[16 points]** Let X and Y be jointly Gaussian random variables with conditional ditribution $f_{Y|X}(v|u) = \sqrt{\frac{1}{2\pi}}e^{-\frac{1}{2}(v-(5-u))^2}$.

   (a) Obtain the best linear estimator $\hat{E}[Y|X = 2]$.

   Solution: Recall that if X and Y are jointly Gaussian, then given $X = u$, $Y \sim N\left(\hat{E}[Y|X = u], MMSE_{\hat{E}[Y|X]}\right)$. Hence, we deduce that $\hat{E}[Y|X = u] = 5 - u$ and hence $\hat{E}[Y|X = 2] = 5 - 2 = 3$.

   (b) Obtain $E[Y|X = 2]$.

   Solution: Recall that if X and Y are jointly Gaussian, then the best unconstrained estimator is equal to the best linear estimator, hence $E[Y|X = 2] = \hat{E}[Y|X = 2] = 3$.

   (c) Obtain $E[Y^2|X = 2]$.

   Solution: Notice that given $X = 2$, $Var(Y|X = 2) = E[Y^2|X = 2] - (E[Y|X = 2])^2$, hence $E[Y^2|X = 2] = Var(Y|X = 2) + (E[Y|X = 2])^2 = MMSE_{\hat{E}[Y|X]} + (E[Y|X = 2])^2 = 1 + (3)^2 = 10$.

5. **[24 points]** Consider a random variable X with pdf:

$$f_X(x; \theta) = \begin{cases} (1 + \theta)x^\theta, & 0 \leq x \leq 1 \\ 0, & \text{else} \end{cases}$$

We want to decide H_1 or H_0 upon the observation of X given the following hypotheses:

$$\begin{aligned} H_0 : &\quad \theta_0 = 2 \\ H_1 : &\quad \theta_1 = 1 \end{aligned}$$

(a) Find the ML decision rule.
Solution: The likelihood ratio test for this problem is given as follows:

$$\Lambda(x) = \frac{f_X(x;\theta_1)}{f_X(x;\theta_0)} \geq \tau$$

for an appropriate threshold τ.

$$\Lambda(x) = \frac{(1+\theta_1)x^{\theta_1}}{(1+\theta_0)x^{\theta_0}} = \frac{2x}{3x^2} = \frac{2}{3x} \geq \tau.$$

Thus, the ML decision rule corresponds to deciding H_1 when

$$x \leq \frac{2}{3},$$

since $\tau = 1$ in this case.

(b) Find the MAP decision rule if the priors $\pi_0 = P(H_0)$ and $\pi_1 = P(H_1)$ satisfy $3\pi_0 = 2\pi_1$.
Solution: As before, the MAP decision rule decides H_1 if

$$x \leq \frac{2}{3}\frac{3}{2} = 1,$$

since $\tau = \frac{\pi_0}{\pi_1} = \frac{2}{3}$ in this case. **Note:** Since $0 \leq x \leq 1$, the MAP decision rule always decides H_1!

(c) Compute $P_{\text{false alarm}}$ for both the ML and MAP decision rules obtained above.
Solution: In the case of a single observation, the corresponding decision rules are: For the ML rule:

$$P_{\text{false alarm}} = P(H_1|H_0) = P\left(X \leq \frac{2}{3}\middle|H_0\right) = \int_0^{2/3} 3x^2dx = \left[x^3\right]_0^{2/3} = \frac{8}{27}.$$

For the MAP rule:

$$P_{\text{false alarm}} = P(H_1|H_0) = 1.$$

(d) Compute $P_{\text{miss detection}}$ for both the ML and MAP decision rules obtained above.
Solution: For the ML rule:

$$P_{\text{miss detection}} = P(H_0|H_1) = P\left\{X > \frac{2}{3}|H_1\right\} = \int_{2/3}^{1} 2xdx = \left[x^2\right]_{2/3}^{1} = 1 - \frac{4}{9} = \frac{5}{9}.$$

For the MAP rule:

$$P_{\text{miss detection}} = P(H_0|H_1) = 0.$$

6. [**8 points**] Assume that a casino player has found a malfunctioning slot machine, which returns an average amount of \$40 every time the player plays. Additionally, the corresponding standard deviation of the amount the machine returns every time is \$5. In order to maximize his revenue, the player has constructed a number of fake slot machine coins so that he can play at no cost. If he plays for 100 times, with what probability will he win at least \$3900, assuming the trials are independent and identically distrbuted.? [Hint: Use Central Limit Theorem]

   Solution: Let X_i denote the amount that the slot machine returns the ith time the player plays. Then, using $n = 100$ fake coins, the player will win $S_n = X_1 + X_2 + \ldots + X_n$ dollars. Therefore,

   $$E[S_n] = 40n, \qquad \mathrm{var}(S_n) = 25n.$$

   By employing the Central Limit Theorem,

   $$\frac{S_n - E[S_n]}{\sqrt{25n}} = \frac{S_n - 4000}{50}.$$

   is approximately distributed as $\mathcal{N}(0,1)$. So

   $$P(S_n \geq 3900) = P\left(\frac{S_n - 4000}{50} \geq \frac{3900 - 4000}{50}\right) \approx Q(-2) = \Phi(2) \approx 0.9772.$$

7. [**28 points**] Emails arrive at Jim's inbox according to a Poisson process with rate $\lambda = 1$ per hour.

   (a) Find the probability that Jim receives at least 3 emails between noon and 2pm.
   Solution: N_2 is Poisson(2). So $P(N_2 \geq 3) = 1 - P(N_2 = 0) - P(N_2 = 1) - P(N_2 = 2) = 1 - e^{-2} - 2e^{-2} - 2^2/2!e^{-2} = 1 - 5e^{-2}$.

   (b) Let T_1 denote the arrival time of the first email. Find the pdf of T_1.
   Solution: T_1 is exponentially distributed with parameter $\lambda = 1$. So $f_{T_1}(t)$ is e^{-t} if $t > 0$ and 0 otherwise.

   (c) Find the mean and variance of T_1.
   Solution: $E[T_1] = \frac{1}{\lambda} = 1$, $\mathrm{Var}(T_1) = \frac{1}{\lambda^2} = 1$.

   (d) Let T_2 denote the arrival time of the second email. Compute the correlation coefficient between T_1 and T_2.

   Solution: Let U denote the interarrival time between the first and second email. Then T_1 and U are iid. Then $\mathrm{Var}(T_2) = \mathrm{Var}(T_1) + \mathrm{Var}(U) = 2\mathrm{Var}(T_1)$ and $\mathrm{Cov}(T_1, T_2) = \mathrm{Cov}(T_1, T_1) = \mathrm{Var}(T_1)$. Therefore $\rho_{T_1,T_2} = \frac{\mathrm{Cov}(T_1,T_2)}{\sqrt{\mathrm{Var}(T_1)\mathrm{Var}(T_2)}} = \frac{1}{\sqrt{2}}$.

   (e) Suppose each email has a chance $\frac{1}{2}$ of being a spam email independently of everything else. Find the probability that Jim receives no spam from noon to 2pm.

Solution:

$$\begin{aligned}
&P(\text{no spam from noon to 2pm})\\
&= \sum_{k=0}^{\infty} P(\text{all } k \text{ emails are not spam}) P(\text{received } k \text{ emails from noon to 2pm})\\
&= \sum_{k=0}^{\infty} 2^{-k} e^{-2} \frac{2^k}{k!} = e^{-1}.
\end{aligned}$$

(f) Let S denote the arrival time of the first spam. Find its CDF and pdf.
Solution: By the same calculation as above, $1 - F_S(t) = P(S > t) = P(\text{no spam in } [0, t]) = \sum_{k=0}^{\infty} 2^{-k} e^{-t} \frac{(t)^k}{k!} = e^{-t/2}$. So S is exponentially distributed with parameter $1/2$, with CDF $F_S(t) = \begin{cases} 1 - e^{-t/2} & t > 0 \\ 0 & t \le 0 \end{cases}$ and pdf $f_S(t) = \begin{cases} \frac{1}{2} e^{-t/2} & t > 0 \\ 0 & t \le 0 \end{cases}$.

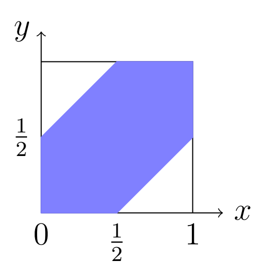
8. **[22 points]** Alice and Bob plan to go to lunch. They each arrive at a random time between noon and 1pm independently of each other. Denote the arrival time of Alice and Bob by X and Y respectively, which are both uniformly distributed over $[0, 1]$ and mutually independent.

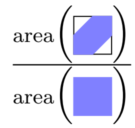
(a) Find the probability that Alice arrives earlier than Bob.
Solution: By symmetry, $P(X < Y) = 1/2$.

(b) Let Z denote the arrival time of the earliest of the two. Find the pdf of Z and the mean $E[Z]$.
Solution: Let $Z = \min(X, Y)$. For any $z \in (0, 1)$, $P(Z \le z) = 1 - P(Z > z) = 1 - P(X > z, Y > z) = 1 - P(X > z)P(Y > z) = 1 - (1 - z)^2$. Therefore the pdf is $f_Z(z) = \begin{cases} 2(1 - z) & 0 < z < 1 \\ 0 & \text{otherwise} \end{cases}$. Then $E[Z] = \int_0^1 2(1 - z) z dz = \frac{1}{3}$.

(c) Let W denote the arrival time of the latest of the two. Find its mean $E[W]$.
Solution: Since $W = \max(X, Y)$ and $X + Y = Z + W$, $E[W] = E[X] + E[Y] - E[Z] = \frac{2}{3}$.

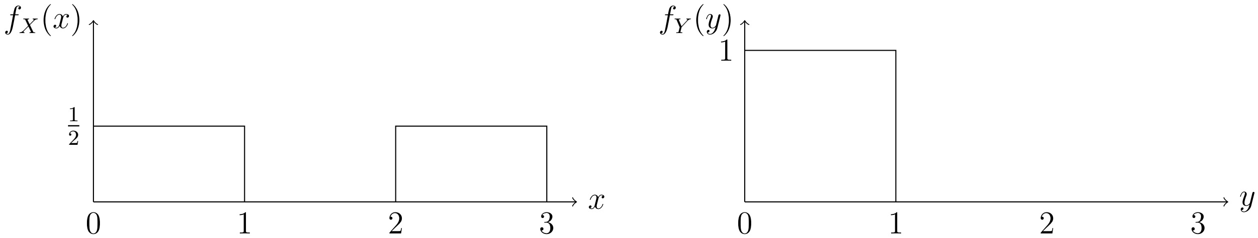
(d) Suppose both Alice and Bob are willing to wait for each other for at most 30 minutes, that is, whoever arrives the first will leave if the other person does not show up in 30 minutes. Find the probability that Alice and Bob meet.
Solution: The region $\{(x, y) : 0 < x, y < 1, |x - y| \le \frac{1}{2}\}$ is drawn as follows.



Then $P(\text{Alice and Bob meet}) = P(|X - Y| \leq 1/2) =$ 

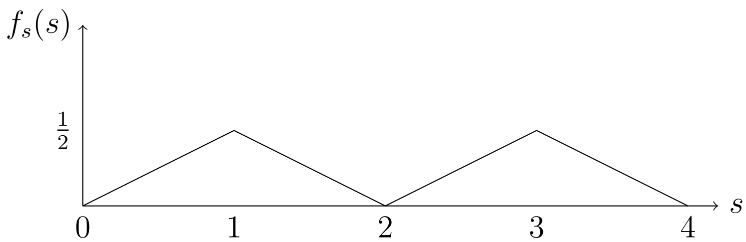
$= \frac{3}{4}$.

9. **[20 points]** Let X and Y be independent random variables with pdfs plotted below:

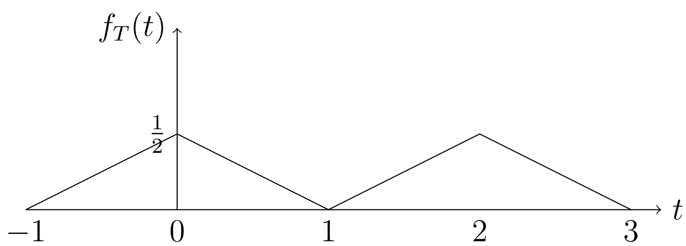


(a) Find the correlation $E[XY]$ and the correlation coefficient $\rho_{X,Y}$.
Solution: By independence $E[XY] = E[X]E[Y] = \frac{3}{4}$ and $\rho_{X,Y} = 0$.

(b) Let $S = X + Y$. Sketch the pdf f_S, clearly labelling all important points.
Solution:



(c) Let $T = X - Y$. Sketch the pdf f_T, clearly labelling all important points.
Solution:



10. **[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 a binary hypothesis testing problem with $H_0 : X \sim \mathcal{N}(0, 1)$ and $H_1 : X \sim \mathcal{N}(1, 1)$.

TRUE	FALSE	
☐	☐	If ML rule is employed, then $P_{\text{false alarm}} = P_{\text{miss detection}}$.
☐	☐	If MAP rule is employed, then $P_{\text{false alarm}}$ and $P_{\text{miss detection}}$ must be different.
☐	☐	$P_e = P_{\text{false alarm}} P(H_1) + P_{\text{miss detection}} P(H_0)$ always holds.

Solution: True, False, False

(b) Let A, B be two events in the sample space Ω with non-zero probabilities. Let $E_1, E_2, \ldots, E_n$ be a partition of Ω.

TRUE	FALSE	
☐	☐	If A, B are mutually exclusive, then A, B must be independent.
☐	☐	$\sum_{i=1}^{n} P(A\|E_i) = \sum_{i=1}^{n} P(B\|E_i)$ always holds.
☐	☐	$\sum_{i=1}^{n} P(E_i\|A) = \sum_{i=1}^{n} P(E_i\|B)$ always holds.

Solution: False, False, True

(c) Assume that $X_1, X_2, \ldots, X_n$ are identically distributed random variables with mean $-\infty < \mu < +\infty$ and variance $0 < \sigma^2 < +\infty$. Let $S_n = X_1 + \ldots + X_n$.

TRUE	FALSE	
☐	☐	If $X_1, X_2, \ldots, X_n$ are also independent, then for any $\epsilon > 0$, $P\left(\left\|\frac{S_n}{n} - \mu\right\| \geq \epsilon\right) \to 0$ as $n \to +\infty$.
☐	☐	If $X_1 = X_2 = \cdots = X_n$, then for any $\epsilon > 0$, $P\left(\left\|\frac{S_n}{n} - \mu\right\| \geq \epsilon\right) \to 0$ as $n \to +\infty$.

Solution: True, False

(d) Flip a biased coin which results in a head with probability 0.55. Define the random variables X and Y as follows: If the outcome is head, set $X = 1$ and $Y = 0$. If the outcome is tail, set $X = 0$ and $Y = 1$. Cross the box to the left of the correct option.

i) The correlation coefficient is $\rho_{X,Y}$ is ☐ 0 ☐ 1 ☐ 0.55 ☐ -1.

ii) The MMSE of estimating Y based on X is: ☐ 1 ☐ 2 ☐ 0.55 ☐ 0 .

Solution: $-1, 0$