18.440 Practice Midterm 2 Partial Solutions

1. (20 points) Let X and Y be independent Poisson random variables with parameter 1. Compute the following. (Give a correct formula involving sums — does not need to be in closed form.)

(a) The probability mass function for X given that $X+Y=5$.

ANSWER: Write $p(k) = P\{X=k\} = e^{-1}/k!$. Then suppose $x \in \{0,1,2,3,4,5\}$. (Mass function is zero at other values.) $P\{X=x|X+Y=5\} = \frac{P\{X=x, X+Y=5\}}{P\{X+Y=5\}} = \frac{P\{X=x\}P\{Y=5-x\}}{P\{X+Y=5\}}$. This is equal to $\frac{p(x)p(5-x)}{\sum_{j=0}^{5} p(x)p(5-x)}$.

(b) The conditional expectation of Y^2 given that $X=2Y$.

ANSWER: Let $y \geq 0$ be an integer. First we compute $P\{Y=y|X=2Y\} = \frac{P\{Y=y, X=2y\}}{P\{X=2Y\}} = \frac{p(2y)p(y)}{\sum_{k=0}^{\infty} p(2k)p(k)}$. Then we note that the $E[Y^2|X=2Y] = \sum_{y=0}^{\infty} P\{Y=y|X=2Y\}y^2$.

(c) The probability mass function for $X-2Y$ given that $X>2Y$.

ANSWER: Write $Z = X-2Y$. Then

$$p_Z(z) = P\{Z=z\} = \sum_{y=-\infty}^{\infty} P\{Y=y\}P\{X=z+2y\} = \sum_{y=0}^{\infty} p(y)p(z+2y).$$

Now for $z>0$ we have $P\{Z=z|Z>0\} = \frac{p_Z(z)}{P\{Z>0\}} = \frac{p_Z(z)}{\sum_{j=1}^{\infty} p_Z(j)}$.

(d) The probability that $X=Y$.

ANSWER: $\sum_{k=0}^{\infty} p(k)^2$.

2. (15 points) Solve the following:

(a) Let X be a normal random variable with parameters (μ, σ^2) and Y an exponential random variable with parameter λ. Write down the probability density function for $X+Y$.

ANSWER:
$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy = \int_0^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-((a-y)-\mu)^2/2\sigma^2} \lambda e^{-\lambda y} dy$

(b) Compute the moment generating function and characteristic function for the uniform random variable on $[0,5]$.

ANSWER: Generally when X is uniform on (a,b) we have $M_X(t) = \frac{e^{tb}-e^{ta}}{t(b-a)}$ and $\phi_X(t) = \frac{e^{itb}-e^{ita}}{it(b-a)}$.

(c) Let $X_1, \ldots, X_n$ be independent exponential random variables of parameter λ. Let Y be the second largest of the X_i. Compute the mean and variance of Y.

ANSWER: This is essentially the radioactive decay problem: how long until $n-1$ of the n particles have decayed. Let T_k be time from when $(k-1)$th particle decays until kth particle decays. Each T_k is exponential with parameter $(n+1-k)\lambda$. It has expectation $\frac{1}{(n+1-k)\lambda}$ and variance $\frac{1}{(n+1-k)^2\lambda^2}$. By additivity of expectation $E[\sum_{k=1}^{n-1} T_k] = \sum_{k=1}^{n-1} E[T_k]$. By independence of the T_k we also have $\mathrm{Var}[\sum_{k=1}^{n-1} T_k] = \sum_{k=1}^{n-1} \mathrm{Var}[T_k]$.

3. (10 points)

(a) Suppose that the pair (X, Y) is uniformly distributed on the disc $x^2 + y^2 \leq 1$. Find f_X, f_Y.

ANSWER: $f(x,y) = \frac{1}{\pi}$ on the disc, zero elsewhere. Then for $x \in (-1,1)$ we have $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \frac{1}{\pi}2\sqrt{1-x^2}$. Symmetric argument gives $f_Y(y) = \frac{1}{\pi}2\sqrt{1-y^2}$ for $y \in (-1,1)$.

(b) Find also $f_{X^2+Y^2}$ and $f_{\max(x,y)}$.

ANSWER: Easiest to first calculate $F_{X^2+Y^2}(a)$ for $a \in [0,1]$. This is $\frac{1}{\pi}$ times the area of the disc of radius r with $r^2 = a$. Thus $F_{X^2+Y^2}(a) = \frac{1}{\pi}\pi r^2 = r^2 = a$. Conclude that $X^2 + Y^2$ is uniform on $[0,1]$ and $f_{X^2+Y^2} = 1$ on $[0,1]$. I'll skip the $f_{\max(x,y)}$ part since it's a bit messy but you could at least in principle compute $F_{\max(x,y)}(a) = P\{\max(x,y) \leq a\}$ using some trigonometry (computing area of intersection of a quadrant with a circle) and then differentiate to get the density function.

(c) Find the conditional probability density for X given $Y = y$ for $y \in [-1,1]$.

ANSWER: Uniform on $(-\sqrt{1-y^2}, -\sqrt{1-y^2})$.

(d) Compute $\mathbb{E}[X^2 + Y^2]$.

ANSWER: $\frac{1}{2}$, by part (b)

4. (10 points) Suppose that X_i are independent random variables which take the values 2 and .5 each with probability 1/2. Let $X = \prod_{i=1}^{n} X_i$.

(a) Compute $\mathbb{E}X$.

ANSWER: By independence $E[X] = \prod_{i=1}^{n} E[X_i] = 1.25^n$.

(b) Estimate the $P\{X > 1000\}$ if $n = 100$.

ANSWER: Let K be number of times X_i is 2, so that $100 - K$ is number of times it is .5. Then
$X = 2^K .5^{100-K} = 2^K/2^{100-K} = 2^{2K-100}$. Note that $X > 1000$ if and only if $2K - 100 \geq 10$, i.e., $K \geq 55$. Now we have a standard binomial problem. What's probability to have at least 55 heads when we toss 100 coins. Standard deviation is $\sqrt{npq} = 5$. So should be roughly $1 - \Phi(1)$.

5. (20 points) Suppose X is an exponential random variable with parameter $\lambda_1 = 1$, Y is an exponential random variable with $\lambda_2 = 2$, and Z is an exponential random variable with parameter $\lambda_3 = 3$. Assume X and Y and Z are independent and compute the following:

(a) The probability density function f_{X+Y}

ANSWER:

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y) dy = \int_0^{\infty} \lambda_1 e^{-\lambda_1 (a-y)} \lambda_2 e^{-\lambda_2 y} dy.$$

(b) $\mathrm{Cov}(XY, X+Y)$

ANSWER: $E[XY(X+Y)] - E[XY]E[X+Y] =$
$E[X^2Y] + E[XY^2] - E[XY](E[X] + E[Y]) =$
$E[X^2]E[Y] + E[X]E[Y^2] - E[X]E[Y]E[X] - E[X]E[Y]E[Y]$. Recall that $E[X] = 1/\lambda_1$ and $E[X^2] = 2/\lambda_2^2$ and $E[Y] = 1/\lambda_2$ and $E[Y^2] = 2/\lambda_2^2$.

(c) $\mathbb{E}[\max\{X, Y, Z\}]$

ANSWER: Perhaps the easiest thing is to first compute

$$F_{\max\{X,Y,Z\}}(a) = P\{\max\{X, Y, Z\} \leq a\} = F_X(a) F_Y(z) F_Z(a) = (1 - e^{\lambda_1 a})(1 - e^{\lambda_2 a})(1 - e^{\lambda_3 a}).$$

Then recall the formula $E[F_{\max\{X,Y,Z\}}] = \int_0^{\infty} 1 - F_{\max\{X,Y,Z\}}(a) da$.

(d) $\mathrm{Var}[\min\{X, Y, Z\}]$

ANSWER: Minimum is exponential with parameter $\lambda_1 + \lambda_2 + \lambda_3 = 6$. So variance is $1/6^2 = 1/36$.

(e) The correlation coefficient $\rho(\min\{X, Y, Z\}, \max\{X, Y, Z\})$.

ANSWER: This one is a bit tricky. Idea is to argue first that $\min\{X, Y, Z\}$ and $\max\{X, Y, Z\} - \min\{X, Y, Z\}$ are independent. Thus by bilinearity of covariance, we have $\mathrm{Cov}(\min\{X, Y, Z\}, \max\{X, Y, Z\}) = \mathrm{Var}\min\{X, Y, Z\} = 1/36$, using the result of (d). From here we use the formula for ρ (though it requires input from (c), which we skipped).

6. (10 points) Suppose $X_1, \ldots, X_{10}$ be independent standard normal random variables. For each $i \in \{2, 3, \ldots, 9\}$ we say that i is a local maximum if $X_i > X_{i+1}$ and $X_i > X_{i-1}$. Let N be the number of local maxima. Compute

(a) The expectation of N.

ANSWER: Each $i \in \{2, 3, \ldots, 9\}$ has a $1/3$ change to be a local maximum. (Basically, X_i has to be the largest among itself and two neighbors.) So $E[N] = 8/3$.

(b) The variance of N.

ANSWER: We need to compute $E[N^2]$. Letting N_i be 1 if i is a local maximum, zero otherwise, we have $E[N^2] = \sum_{i=2}^{9} \sum_{j=2}^{9} E[N_i N_j]$. This sum includes

(a) 8 terms of form $E[N_i N_i]$, which are each equal to $1/3$.

(b) 14 terms of form $E[N_i N_{i+1}]$ or $E[N_i N_{i-1}]$ which are each equal to zero (to neighbors can't both be local maxima).

(c) 12 terms of form $E[N_i N_{i+2}]$ or $E[N_i N_{i-2}]$. Consider: five guys in a row, we need the probability that the second and fourth are both local maxima. Have $1/5$ chance that second is largest among these five, and given that, have $1/3$ chance that fourth is local maximum. So $1/15$ chance in case second is largest, similarly $1/15$ in case fourth is largest, so $E[N_i N_{i+2}] = 2/15$ over all.

(d) $64 - 8 - 14 - 12 = 30$ remaining terms.

Sum all these up to get $E[N^2]$. Then $\mathrm{Var}[N^2] = E[N^2] - E[N]^2$.

(c) The correlation coefficient $\rho(N, X_1)$.

ANSWER: Here $\rho(N, X_1) = \text{Cov}(N, X_1)/\sqrt{\text{Var}(N)\text{Var}(X_1)}$. Hard part remaining is to compute $\text{Cov}(N, X_1)$. By bilinearity of expectation, this is $\sum_{j=2}^{9} \text{Cov}(N_j, X_1)$. These terms are all zero except when $j = 2$. So we just need to find $\text{Cov}(N_2, X_1)$. Key step to compute $E[N_2 X_1]$. If f is standard normal density, this can be written as $\int\int\int f(x)f(y)f(z)x$ where the integral is taken over the portion of R^2 for which $y > x$ and $y > z$. I think I'll skip the part where we actually compute the integral.

7. (15 points) Give the name and an explicit formula for the density or mass function of $\sum_{i=1}^{n} X_i$ when the X_i are

(a) Independent normal with parameter μ, σ^2.

ANSWER: normal with parameters $(n\mu, n\sigma^2)$

(b) Independent exponential with parameter λ.

ANSWER: gamma with parameters n and λ

(c) Independent geometric with parameter p.

ANSWER: negative binomial with parameter p.

(d) Independent Poisson with parameter λ

ANSWER: Poisson with parameter $n\lambda$.

(e) Independent Bernoulli with parameter p.

ANSWER: binomial with parameters (n, p).