MAST90105 Methods of Mathematical Statistics
Assignment 2, Semester 1, 2019 **Solutions**

1. A low pressure system (low) that forms in the Australian region has a chance of 8% to transform into a tropical cyclone.

   (a) Assume that 20 lows will form in April 2019. Find the probability that two of them will transform into a tropical cyclone.

   - *Let X be the number lows that will transform into a tropical cycolne in April 2019. Then X has a binomial distribution with $p = 0.08, n = 20$ and*
   - $\Pr(X = 2) = \binom{20}{2} p^2(1-p)^{18} = 0.271.$
   - *To compute the probability using R:* `dbinom(2,20,0.08)`

   (b) Find the probability that the first low that will transform into a tropical cyclone in November 2019 will be the 5th low that will form in this month.

   - *Let N_1 be the number of lows that form in November 2019 when one of them transforms into a tropical cycolne for the first time. Then N_1 has a geometric distribution with $p = 0.08$ and*
   - $\Pr(N_1 = 5) = p(1-p)^4 = 0.057.$
   - *To compute this probability using R:* `dgeom(4,0.08)`

   (c) Find the probability that the second low that will transform into a tropical cyclone in December 2019 will be the 15th low that will form in this month.

   - *Let N_2 be the number of lows that form in December 2019 when one of them transforms into a tropical cycolne for the second time. Then N_2 has a negative binomial distribution with $p = 0.08, r = 2$ and*
   - $\Pr(N_2 = 15) = \binom{14}{1} p^2(1-p)^{13} = 0.03.$
   - *To compute this probability using R:* `dnbinom(13,2,0.08)`

   (d) How many lows are expected to form before two of them transform to tropical cyclones?

   - $\mathrm{E}(N_2) = r/p = 25.$

   (e) Assume that 150 lows in total will form between November 1st, 2019 and April 30th, 2020. Find the probability that at least 10 of them will transform into tropical cyclones.

   - Compute the exact probability
     - *Let T be the number of lows that will transform into tropical cyclones. Then T has a binomial distribution with parmeters $p = 0.08, n = 150$ and*
     - $\Pr(T \geq 10) = 1 - \Pr(T \leq 9) = 1 - \sum_{k=0}^{9} \binom{150}{k} p^k(1-p)^{n-k} = 0.769.$
     - *To compute this probability using R:* `1−pbinom(9,150,0.08)`
   - Compute this probability using the Poisson approximation

- *The approximate distribution of T is a Poisson distribution with parameter $\lambda = np = 12$ and therefore*
- $\Pr(T \geq 10) = 1 - \Pr(T \leq 9) \approx 1 - \sum_{k=0}^{9} e^{-\lambda}\frac{\lambda^k}{k!} = 0.758.$
- *To compute this probability using R:* `1-ppois(9,12)`

2. The telephone calls received by a call center follow a Poisson process with the rate of 8 calls per minute.

(a) Find the probability that the call center receives exactly 45 calls in 5 minutes.

- *Let X_5 be the number of telephone calls received by the call center in 5 minutes. Then X has a Poisson distribution with parameter $\lambda = 5 \cdot 8 = 40$ and*
- $\Pr(X_5 = 45) = e^{-\lambda}\frac{\lambda^{45}}{45!} = 0.044.$
- *To compute this probability using R:* `dpois(45,40)`

(b) Find the probability that the call center receives at least 500 calls in one hour and at most 500 calls in the following hour.

- *Let Y_1 and Y_2 be the number of calls received by the call center in the first and second hours, respectively. These are independent random variables and they follow a Poisson distribution with parameter $\lambda = 60 \cdot 8 = 480$. We find:*
- $\Pr(Y_1 \geq 500) = 1 - \Pr(Y_1 \leq 499) = 1 - \sum_{k=0}^{499} e^{-\lambda}\frac{\lambda^k}{k!} = 0.186,$
- $\Pr(Y_2 \leq 500) = \sum_{k=0}^{500} e^{-\lambda}\frac{\lambda^k}{k!} = 0.826,$
- $\Pr(Y_1 \geq 500, Y_2 \leq 500) = \Pr(Y_1 \geq 500) \cdot \Pr(Y_2 \leq 500) = 0.154$
- *To compute this probability using R:* `(1-ppois(499,480))*ppois(500,480)`

(c) How many calls the call center is expected to receive in 8 hours? Find the standard deviation of the number of calls received in 8 hours.

- *The expected number of calls in 8 hours is* $8 \cdot 8 \cdot 60 = 3840$
- *The standard deviation is* $\sqrt{3840} \approx 62$

(d) Assume now that the call center cannot receive more than 10 calls in 1 minute due to limited capacity. If more calls arrive, then some of them will not be answered. Find the probability that all calls will be answered in 10 minutes. Find the mean of the time (in minutes) to the first unanswered call.

- *Let T_i be the number of calls received by the call center in the ith minute. $T_1, \ldots, T_{10}$ are independent and follow a Possion distribution with parameter $\lambda = 8$ and therefore*
- $\Pr(T_i \leq 10) = \sum_{k=0}^{10} e^{-\lambda}\frac{\lambda^k}{k!} = 0.816,$
- $\Pr(T_1 \leq 10, \ldots, T_{10} \leq 10) = \{\Pr(T_1 \leq 10)\}^{10} = 0.131.$
- *To compute this probability using R:* `ppois(10,8)^10`
- *Let T_0 be the time (in minutes) to the first unanswered call. T_0 follows a geometric distribution with parameter $p = 1 - 0.816 = 0.184$ and*
- $\mathrm{E}(T_0) = 1/p = 5.43.$

3. A random variable X has a pdf given by

$$f(x) = \begin{cases} Ce^{2x}, & x < 0, \\ Ce^{-x}, & x \geq 0. \end{cases}$$

(a) Find the normalizing constant C.

- $\int_{-\infty}^{\infty} f(x)dx = C\int_{-\infty}^{0} e^{2x}dx + C\int_{0}^{\infty} e^{-x}dx = \frac{C}{2}e^{2x}\Big|_{-\infty}^{0} - Ce^{-x}\Big|_{0}^{\infty} = \frac{3C}{2}$
- $\int_{-\infty}^{\infty} f(x)dx = 1$ and therefore $C = \frac{2}{3}$.

(b) Find the cdf of X.

- $F(x) = \begin{cases} C\int_{-\infty}^{x} e^{2t}dt = \frac{C}{2}e^{2t}\Big|_{-\infty}^{x} = \frac{1}{3}e^{2x}, & x < 0, \\ C\int_{-\infty}^{0} e^{2t}dt + C\int_{0}^{x} e^{-t}dt = \frac{C}{2}e^{2t}\Big|_{-\infty}^{0} - Ce^{-t}\Big|_{0}^{x} = 1 - \frac{2}{3}e^{-x}, & x \geq 0. \end{cases}$

(c) Let $-2 < t < 1$. Find $M_X(t) = \mathrm{E}(e^{Xt})$, the moment generating function of X.

- $M_X(t) = C\int_{-\infty}^{0} e^{tx}\cdot e^{2x}dx + C\int_{0}^{\infty} e^{tx}\cdot e^{-x}dx = \frac{C}{2+t}e^{(2+t)x}\Big|_{-\infty}^{0} - \frac{C}{1-t}e^{-(1-t)x}\Big|_{0}^{\infty} = \frac{C}{2+t} + \frac{C}{1-t} = \frac{2}{(2+t)(1-t)}$.

(d) Use $M_X(t)$ to find the mean and variance of X, $\mathrm{E}(X)$ and $\mathrm{Var}(X)$.

- $\mathrm{E}(X) = M_X'(t)\Big|_{t=0} = \left(-\frac{C}{(2+t)^2} + \frac{C}{(1-t)^2}\right)\Big|_{t=0} = -\frac{C}{4} + C = \frac{1}{2}$,
- $\mathrm{E}(X^2) = M_X''(t)\Big|_{t=0} = \left(\frac{2C}{(2+t)^3} + \frac{2C}{(1-t)^3}\right)\Big|_{t=0} = \frac{C}{4} + 2C = \frac{3}{2}$,
- $\mathrm{Var}(X) = \mathrm{E}(X^2) - \{\mathrm{E}(X)\}^2 = \frac{3}{2} - \frac{1}{4} = \frac{5}{4}$.

(e) Find the median of X.

- *Let m_X be the median of X. Since $F(0) = 1/3$, we have $m_X > 0$. We find:*
- $F(m_X) = 1 - \frac{2}{3}e^{-m_X} = \frac{1}{2} \Rightarrow m_X = \ln 4 - \ln 3 = 0.288$.

4. Let X and Y have the joint pdf defined by $f(x,y) = \begin{cases} C(x+y), & 0 < x < 1, 0 < y < 1, \\ 0, & \text{otherwise.} \end{cases}$

(a) Find the normalizing constant C.

We have: $\int_0^1\int_0^1 f(x,y)dxdy = C\int_0^1\int_0^1 xdxdy + C\int_0^1\int_0^1 ydxdy = \frac{C}{2} + \frac{C}{2} = C = 1$.

(b) Find $F(x,y) = \Pr(X < x, Y < y)$ for any x, y and compute $\Pr(X < 0.4, Y < 0.6)$.

- *If* $0 < x < 1, 0 < y < 1$: $F(x,y) = \int_0^x\int_0^y f(x^*,y^*)dx^*dy^* = \int_0^x\int_0^y x^*dx^*dy^* + \int_0^x\int_0^y y^*dx^*dy^* = 0.5(x^2y + xy^2) = 0.5xy(x+y)$,
- *If* $0 < x < 1, y > 1$: $F(x,y) = \int_0^x\int_0^1 f(x^*,y^*)dx^*dy^* = \int_0^x\int_0^1 x^*dx^*dy^* + \int_0^x\int_0^1 y^*dx^*dy^* = 0.5(x^2 + x) = 0.5x(x+1)$,
- *If* $x > 1, 0 < y < 1$: $F(x,y) = 0.5y(y+1)$,

- *If* $x < 0$ *or* $y < 0$: $F(x, y) = 0$.
- $\Pr(X < 0.4, Y < 0.6) = F(0.4, 0.6) = 0.12$.

(c) Compute $\mu_X = \mathrm{E}(X)$, $\mu_Y = \mathrm{E}(Y)$, $\sigma_X^2 = \mathrm{Var}(X)$, $\sigma_Y^2 = \mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$, and $\rho = \mathrm{Cor}(X, Y)$.

- *The marginal cdf and pdf of* X *is* $F_X(x) = F(x, 1) = 0.5x(x+1)$ *and* $f_X(x) = F_X'(x) = x + 0.5$ *for* $0 < x < 1$,
- $\mu_X = \int_0^1 x f_X(x)dx = \int_0^1 (x^2 + 0.5x)dx = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$,
- $\mathrm{E}(X^2) = \int_0^1 x^2 f_X(x)dx = \int_0^1 (x^3 + 0.5x^2)dx = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}$,
- $\mathrm{Var(X)} = \mathrm{E}(X^2) - \mu_X^2 = \frac{5}{12} - \frac{49}{144} = \frac{11}{144}$,
- *Similarly,* $\mu_Y = \frac{7}{12}$ *and* $\mathrm{Var(Y)} = \frac{11}{144}$,
- $\mathrm{E}(XY) = \int_0^1 \int_0^1 xyf(x, y)dxdy = \int_0^1 \int_0^1 xy(x + y)dxdy = \int_0^1 x^2dx \int_0^1 ydy + \int_0^1 y^2dy \int_0^1 xdx = \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{3}$,
- $\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mu_X\mu_Y = \frac{1}{3} - \frac{49}{144} = -\frac{1}{144}$,
- $\rho = \mathrm{Cov}(X, Y)/(\sigma_X\sigma_Y) = -\frac{1}{11}$.

5. Let X and Y be jointly normal random variables with parameters $\mathrm{E}(X) = \mathrm{E}(Y) = 0$, $\mathrm{Var}(X) = \mathrm{Var}(Y) = 1$, and $\mathrm{Cor}(X, Y) = \rho$.

(a) Find ρ such that $W_1 = 0.6X + 0.8Y$ and $W_2 = 0.8X + 0.6Y$ are independent.

- W_1 *and* W_2 *are normal random variables and they are independent if and only if* $\mathrm{Cov}(X, Y) = 0$
- $\mathrm{Cov}(W_1, W_2) = 0.6 \cdot 0.8\mathrm{Cov}(X, X) + (0.6 \cdot 0.6 + 0.8 \cdot 0.8)\mathrm{Cov}(X, Y) + 0.8 \cdot 0.6\mathrm{Cov}(Y, Y) = 0.48 + (0.36 + 0.64)\rho + 0.48 = 0.96 + \rho = 0 \Rightarrow \rho = -0.96$.

(b) Find $\Pr(X < 1|X = Y)$ if $\rho = 0.6$.

- X *and* $W = X - Y$ *are jointly normal and* $\mathrm{Cov}(X, W) = \mathrm{E}(XW) = \mathrm{E}(X^2) - \mathrm{E}(XY) = 1 - \rho$ *because* $\mathrm{E}(X) = 0$ *and* $\mathrm{E}(XY) = \mathrm{Cov}(X, Y) + \mathrm{E}(X)\mathrm{E}(Y) = \rho$
- $\mathrm{Var(W)} = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{Cov}(X, Y) = 2 - 2\rho$
- $\rho^* = \mathrm{Cor}(X, W) = \mathrm{Cov}(X, W)/\sqrt{\mathrm{Var(X)Var(W)}} = \sqrt{(1 - \rho)/2}$
- *The conditional distribution of* $X|W = 0$ *is normal with mean* $\rho^* \cdot 0 = 0$ *and variance* $1 - (\rho^*)^2 = (1 + \rho)/2 = 0.8$ *and therefore* $\Pr(X < 1|W = 0) = 0.868$.
- *To compute the probability using R:* `pnorm(1,mean=0,sd=sqrt(0.8))`