MATH562 – Fall 2025

Problem Set: Week 1

1. (a) Suppose we draw a sample of two balls at random without replacement from a bag containing two red balls and three white balls. Give the sample space for this random experiment. Are its elements equiprobable? What is the probability that both balls in the sample are red, given that at least one is red? **Solution**: The sample space is $\Omega = \{\{R,R\},\{R,W\},\{W,W\}\}$, since there is no mention of order in the sampling. The elements of Ω are not equiprobable:

$$\Pr\{\{R,R\}\} = \frac{\binom{3}{0}\binom{2}{2}}{\binom{5}{2}} = \frac{1}{10}, \quad \Pr\{\{R,W\}\} = \frac{\binom{3}{1}\binom{2}{1}}{\binom{5}{2}} = \frac{6}{10}, \quad \Pr\{\{R,R\}\} = \frac{\binom{3}{2}\binom{2}{0}}{\binom{5}{2}} = \frac{3}{10}.$$

Let A and B denote the events 'both balls red' and 'at least one red' respectively. So, $A = \{\{R,R\}\}$ and $B = \{\{R,W\},\{R,R\}\}$, and since $A \subset B$,

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A)}{\Pr(B)} = \frac{1/10}{1/10 + 6/10} = \frac{1}{7}.$$

(b) Find the median of $Y = \exp(X)$, where $X \sim \mathrm{Unif}(a,b)$ for $a < b$. **Solution**: Let $y_{0.5}$ be the median. Then as $X \sim U(a,b)$,

$$\Pr(Y \le y_{0.5}) = \Pr(X \le \log y_{0.5}) = \Pr\{X \le (a+b)/2\},$$

so $y_{0.5} = \exp\{(a+b)/2\}$.

(c) The joint probability mass function of random variables (X,Y) is given by the table:

	x		
y	1	3	5
2	c	$2c$	$3c$
4	$3c$	$2c$	c

Find $\mathrm{E}(X)$ and $\mathrm{E}(X \mid Y=4)$. Are X and Y independent? **Solution**:
Since all the entries of the table must sum to 1, we must have $c = 1/12$. Hence

$$\mathrm{E}(X) = 1 \cdot \frac{4}{12} + 3 \cdot \frac{4}{12} + 5 \cdot \frac{4}{12} = \frac{36}{12} = 3,$$

and the conditional expectation is

$$\mathrm{E}(X \mid Y=4) = \frac{1 \times 3/12 + 3 \times 2/12 + 5 \times 1/12}{6/12} = \frac{14}{6}.$$

The random variables X and Y are clearly dependent, since $\mathrm{E}(X \mid Y=4) \ne \mathrm{E}(X)$.

(d) A simple model for a daily rainfall amount X is that $X = 0$ with probability $1-p$ and otherwise is exponential with parameter λ. Find the probability that $X = 0$ given that $X < 3$, and obtain the mean and variance of X.

Solution:

We have

$$\Pr(X < 3) = \Pr(X = 0) + \Pr(0 < X < 3 \mid X > 0)\Pr(X > 0) = (1-p) + p(1 - e^{-3\lambda});$$

equivalently $\Pr(X < 3) = 1 - \Pr(X \geq 3) = 1 - pe^{-3\lambda}$. Hence $\Pr(X = 0 \mid X < 3) = (1-p)/(1-pe^{-3\lambda})$. The mean and variance of X, p/λ and $(2-p)p/\lambda^2$, can be computed by noting that

$$\mathrm{E}(X^r) = \mathrm{E}(X^r \mid X = 0)\Pr(X = 0) + \mathrm{E}(X^r \mid X > 0)\Pr(X > 0) = 0^r(1-p) + \lambda^{-r}\Gamma(r+1)p,$$

or via the moment-generating function,

$$M_X(t) = \mathrm{E}(e^{tX}) = \mathrm{E}(e^{tX} \mid X = 0)\Pr(X = 0) + \mathrm{E}(e^{tX} \mid X > 0)\Pr(X > 0) = 1 - p + p\frac{\lambda}{\lambda - t}, \quad t < \lambda.$$

(e) If $\Pr(X > x) = 1/x^2$, for $x > 1$, find the probability density function of $Y = 1/X$. **Solution**: For $y \in (0,1]$, $F_Y(y) = \Pr(Y \leq y) = \Pr(1/X \leq y) = \Pr(X \geq 1/y) = y^2$. So,

$$f_Y(y) = \frac{\partial F_Y(y)}{\partial y} = 2y, \quad 0 < y \leq 1.$$

(f) If $X \sim \exp(\lambda)$, find the distribution and density functions of $Y = \cos X$. **Solution**: $Y = \cos X$ takes values only in the range $-1 \leq y \leq 1$, so

$$F_Y(y) = \begin{cases} 0, & y < -1 \\ 1, & y \geq 1. \end{cases}$$

A sketch of $\cos x$ for $x \geq 0$ shows that in the range $0 < x < 2\pi$, and for $-1 < y < 1$, we have $\cos X \leq y \Leftrightarrow \cos^{-1} y \leq X \leq 2\pi - \cos^{-1} y$. Since the cosine function is periodic, we therefore have

$$\cos X \leq y \quad \Leftrightarrow \quad X \in \bigcup_{j=0}^{\infty} \{x : 2\pi j + \cos^{-1} y \leq x \leq 2\pi(j+1) - \cos^{-1} y\},$$

and thus

$$\begin{aligned} \Pr(Y \leq y) &= \sum_{j=0}^{\infty} \Pr\left\{2\pi j + \cos^{-1} y \leq X \leq 2\pi(j+1) - \cos^{-1} y\right\} \\ &= \sum_{j=0}^{\infty} \left(\exp[-\lambda\{2\pi j + \cos^{-1} y\}] - \exp[-\lambda\{2\pi(j+1) - \cos^{-1} y\}]\right) \\ &= \frac{\exp(-\lambda\cos^{-1} y) - \exp(\lambda\cos^{-1} y - 2\pi\lambda)}{1 - \exp(-2\pi\lambda)}, \end{aligned}$$

where we noticed that the summation is proportional to a geometric series.
If $y = 1$, then $\cos^{-1} y = 0$, and so $\Pr(Y \leq 1) = 1$, and if $y = -1$, then $\cos^{-1} y = \pi$, and then $\Pr(Y \leq -1) = 0$, as required. Here we used values of $\cos^{-1} y$ in the range $[0, \pi]$.
The density function is found by differentiation: since $\cos(\cos^{-1} y) = y$, we have

$$\frac{d\cos^{-1} y}{dy} = -\frac{1}{\sin(\cos^{-1} y)},$$

and this gives

$$f_Y(y) = \frac{\lambda}{\sin(\cos^{-1} y)} \times \frac{\exp(-\lambda\cos^{-1} y) + \exp(\lambda\cos^{-1} y - 2\pi\lambda)}{1 - \exp(-2\pi\lambda)}, \quad y \in (-1, 1).$$

(g) Two successive software downloads take times (minutes) $X_1 \sim \mathcal{N}(8, 3^2)$ and $X_2 \sim \mathcal{N}(16, 4^2)$. The download times are independent. Find the following:

i. the distribution of the total download time $T = X_1 + X_2$,

ii. the probability that T exceeds 30 minutes,

iii. the probability that T exceeds 30 minutes, given that $X_1 = 10$, and

iv. the probability that X_1 was less than 7 minutes, given that $T = 30$ minutes.

Solution: (i) The random variable $T = X_1 + X_2$ follows a normal distribution since a linear combination of normal variables is normal, and $\mathrm{E}(T) = \mathrm{E}(X_1 + X_2) = \mathrm{E}(X_1) + \mathrm{E}(X_2) = 8 + 16 = 24$ and $\mathrm{Var}(T) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) = 9 + 16 = 25$, so $T \sim \mathcal{N}(24, 5^2)$.

(ii) The random variable $Z = (T - 24)/5 \sim \mathcal{N}(0, 1)$, so

$$\Pr(T > 30) = \Pr\left(Z > \frac{30 - 24}{5}\right) = 1 - \Pr(Z \leq 1.2) = 1 - \Phi(1.2) = 1 - 0.88493 \approx 0.115,$$

where $\Phi(\cdot)$ denotes the standard normal CDF.

(iii) The probability that the total download time T exceeds 30 minutes given that $X_1 = 10$ is

$$\Pr(T > 30 \mid X_1 = 10) = \Pr(X_1 + X_2 > 30 \mid X_1 = 10) = \Pr(X_2 > 20 \mid X_1 = 10) = \Pr(X_2 > 20)$$

by the independence of X_1 and X_2. The random variable $Z_2 = (X_2 - 16)/4$ follows a standard normal distribution, so

$$\Pr(X_2 > 20) = 1 - \Pr\left(Z_2 \leq \frac{20 - 16}{4}\right) = 1 - \Phi(1) = 1 - 0.84134 \approx 0.152.$$

(iv) $Y = (X_1, T)^\top = (X_1, X_1 + X_2)^\top$ is a linear combination of normal variables, so it has a bivariate normal distribution, with mean and covariance matrix

$$\mu = \begin{pmatrix} \mathrm{E}(X_1) \\ \mathrm{E}(T) \end{pmatrix} = \begin{pmatrix} 8 \\ 24 \end{pmatrix}, \quad \Omega = \begin{pmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, T) \\ \mathrm{Cov}(X_1, T) & \mathrm{Var}(T) \end{pmatrix} = \begin{pmatrix} 9 & 9 \\ 9 & 25 \end{pmatrix},$$

since $\mathrm{Cov}(X_1, T) = \mathrm{Var}(X_1) = 9$ by the independence of X_1 and X_2. Thus

$$\begin{pmatrix} X_1 \\ T \end{pmatrix} \sim \mathcal{N}_2 \left\{ \begin{pmatrix} 8 \\ 24 \end{pmatrix}, \begin{pmatrix} 9 & 9 \\ 9 & 25 \end{pmatrix} \right\}.$$

Now $X_1 \mid T = 30 \sim \mathcal{N}(\mu, \sigma^2)$ where $\mu = 8 + 9 \times (30 - 24)/25 = 10.16$ and $\sigma^2 = 9 - 9^2/25 = 2.4^2$. Thus, $Z_1 = (X_1 - \mu)/\sigma \mid T = 30 \sim \mathcal{N}(0, 1)$, so

$$\Pr(X_1 < 7 \mid T = 30) = \Pr\left(Z_1 < \frac{7 - 10.16}{2.4}\right) \approx \Phi(-1.32) = 1 - \Phi(1.32) = 1 - 0.90658 \approx 0.093.$$

(h) If $Z_j \overset{i.i.d.}{\sim} \mathcal{N}(a_j, 1)$ $(j = 1, \ldots, n)$ independent of $Y \sim \mathcal{N}(\mu, \sigma^2)$, find the joint distribution of $X_j = Y + Z_j$ and state under what circumstances they are finitely exchangeable. **Solution**: We can write

$$X_{n\times 1} = (X_1, \ldots, X_n)^\top = (Y + Z_1, \ldots, Y + Z_n)^\top = B_{n\times(n+1)}(Y, Z_1, \ldots, Z_n)^\top,$$

where $(Y, Z_1, \ldots, Z_n)^\top \sim N_{n+1}\{(\mu, a_1, \ldots, a_n)^\top, \mathrm{diag}(\sigma^2, 1, \ldots, 1)\}$. Hence X has a joint normal distribution (see slide 20), and it is easy to check that $\mathrm{E}(X) = (\mu + a_1, \ldots, \mu + a_n)^\top$ and $\mathrm{Var}(X) = I_n + \sigma^2 1_n 1_n^\top$. This distribution is finitely exchangeable if we have constant $\mathrm{Var}(X_j)$, constant $\mathrm{Cov}(X_j, X_k)$ for $j \neq k$, which are both true, and equal means; the latter occurs if and only if all the a_j are equal.

(i) Three friends arrive to dine together independently at random between 7 and 8 o'clock. Give the density functions of the times of the first and last arrivals, and the expected time between them. **Solution**: Let X_1, X_2, X_3 denote the arrival times after 19.00; as they arrive independently at random, we can suppose that $X_1, X_2, X_3 \overset{i.i.d.}{\sim} U(0,1)$. The arrival times of the first and the last are $U = \min(X_1, X_2, X_3)$ and $V = \max(X_1, X_2, X_3)$. The densities can be obtained directly or ...

As the X_j have distribution function $F(u) = u$ in $0 < u < 1$, we obtain

$$\begin{aligned}
\Pr(U \le u) &= 1 - \Pr(X_1 > u, X_2 > u, X_3 > u) \\
&= 1 - \Pr(X_1 > u)\Pr(X_2 > u)\Pr(X_3 > u) \\
&= 1 - (1-u)^3, \quad 0 < u < 1,
\end{aligned}$$

and the density is $f_U(u) = 3(1-u)^2$, for $0 < u < 1$. Likewise

$$\Pr(V \le v) = \Pr(X_1 \le v)\Pr(X_2 \le v)\Pr(X_3 \le v) = v^3, \quad 0 < v < 1,$$

and the density is $f_V(v) = 3v^2$, for $0 < v < 1$.

For the final part, we seek $\mathrm{E}(V - U) = \mathrm{E}(V) - \mathrm{E}(U)$, and

$$\mathrm{E}(V) = \int_0^1 v f_V(v)\,\mathrm{d}v = \frac{3}{4}, \quad \mathrm{E}(U) = \int_0^1 u f_u(u)\,\mathrm{d}u = \int_0^1 3u(1-u)^2\,\mathrm{d}u = \frac{1}{4},$$

using integration by parts. Thus $\mathrm{E}(V - U) = 2/4$, or 30 minutes.

(j) Recall that the moment-generating function of a random variable X is $M_X(t) = \mathrm{E}(e^{tX})$, defined for all $t \in \mathbb{R}$ such that $M_X(t) < \infty$. If $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \exp(\lambda)$ and $S = \sum_{j=1}^n X_j$, find $M_S(t)$. **Solution**: The density of X is $f(x) = \lambda\exp(-\lambda x)$, for $x > 0$ and $\lambda > 0$, so

$$M_X(t) = \mathrm{E}\left\{\exp(tX)\right\} = \int_0^\infty \exp(tx)\lambda\exp(-\lambda x)\mathrm{d}x = \frac{\lambda}{\lambda - t}, \quad t < \lambda.$$

Hence independence of $X_1, \ldots, X_n$ gives

$$M_S(t) = \mathrm{E}\left\{\exp(tS)\right\} = \mathrm{E}\left\{\exp(tX_1 + \cdots + tX_n)\right\} = \prod_{j=1}^n \mathrm{E}\left\{\exp(tX_j)\right\} = \frac{\lambda^n}{(\lambda - t)^n}, \quad t < \lambda.$$

(k) If $X_1, \ldots, X_N \overset{i.i.d.}{\sim} \mathrm{Poiss}(\lambda)$, $\Pr(N = n) = p(1-p)^{n-1}$ for $n \in \{1, 2 \ldots\}$, and $\mathbf{1}(\cdot)$ is the indicator function, find the mean and variance of $T = \sum_{j=1}^N \mathbf{1}(X_j = 0)$. How would the mean of T change if $X_1, \ldots, X_N$ were dependent? **Solution**: Since $X_1, \ldots, X_N \overset{i.i.d.}{\sim} \mathrm{Poiss}(\lambda)$, each indicator variable $\mathbf{1}(X_j = 0)$ is Bernoulli with success probability $q = e^{-\lambda}$. So, conditional on $N = n$, $T \sim \mathrm{Bin}(n, q)$, which gives $\mathrm{E}(T \mid N = n) = nq$ and $\mathrm{Var}(T \mid N = n) = nq(1-q)$. As N has a geometric distribution with success probability p, we have $\mathrm{E}(N) = 1/p$ and $\mathrm{Var}(N) = (1-p)/p^2$, and thus

$$\begin{aligned}
\mathrm{E}(T) &= \mathrm{E}_N\,\mathrm{E}(T \mid N) = \mathrm{E}_N(Nq) = q/p, \\
\mathrm{Var}(T) &= \mathrm{E}_N\,\mathrm{Var}(T \mid N) + \mathrm{Var}_N\,\mathrm{E}(T \mid N) \\
&= \mathrm{E}_N\{Nq(1-q)\} + \mathrm{Var}_N(Nq) = q(1-q)/p + q^2(1-p)/p^2.
\end{aligned}$$

If the X_j's are dependent, $\mathrm{E}(T)$ is unchanged because expectation is a linear operator.

(l) If $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \mathrm{Pois}(\lambda)$, find the approximate distribution of $Y = 2\bar{X}^{1/2}$ for large n. **Solution**: The central limit theorem implies that $\bar{X} \sim \mathcal{N}(\mu, \mu/n)$, so applying the delta method with $g(u) = 2\sqrt{u}$, giving $g'(u) = u^{-1/2}$, leads to

$$Y = g(\bar{X}) \sim \mathbb{N}\left\{g(\mu), g'(\mu)^2 \times \mu/n\right\} = \mathcal{N}(2\sqrt{\mu}, 1/n), \quad n \to \infty.$$

Thus the variance of Y does not depend on μ, at least to this order of approximation: the square root transformation is variance-stabilizing for the Poisson distribution. The related *Anscombe transformation* $Y = (4\bar{X} + 3/2)^{1/2}$ is widely used in certain imaging settings.

(m) If $(X_1, Y_n), \ldots, (X_n, Y_n)$ form a random sample from a bivariate distribution with means and (finite positive) variances (μ_X, μ_Y) and (σ_X^2, σ_Y^2) and correlation ρ, with $\mu_X \neq 0$, show that $T = \sum_j Y_j / \sum_j X_j$ converges in probability to $\theta = \mu_Y/\mu_X$ and that for large n

$$n^{1/2}\mu_X(T - \theta) \sim \mathcal{N}(0, \sigma_X^2\theta^2 - 2\theta\rho\sigma_X\sigma_Y + \sigma_Y^2).$$

Provide an estimator for the variance.

Solution: As $n \to \infty$, the weak law of large numbers gives (in the usual notation) $\bar{X} \xrightarrow{\mathbb{P}} \mu_X$ and $\bar{Y} \xrightarrow{\mathbb{P}} \mu_Y$, so as these convergences are also in distribution, Slutsky's theorem gives $T \rightsquigarrow \mu_Y/\mu_X = \theta$. As this is convergence in distribution to a constant, we also have $T \xrightarrow{\mathbb{P}} \theta$. The (multivariate) central limit theorem gives

$$n^{1/2}\left\{\begin{pmatrix}\bar{X}\\ \bar{Y}\end{pmatrix} - \begin{pmatrix}\mu_X\\ \mu_Y\end{pmatrix}\right\} \rightsquigarrow \mathbb{N}\left\{\begin{pmatrix}0\\ 0\end{pmatrix}, \begin{pmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix}\right\}, \quad n \to \infty,$$

and as $\theta = g(\mu_X, \mu_Y) = \mu_Y/\mu_X$ has derivatives $g_1(\mu_X, \mu_Y) = -\mu_Y/\mu_X^2 = -\theta/\mu_X$ and $g_2(\mu_X, \mu_Y) = 1/\mu_X$, simplifying the variance expression

$$\sigma_\theta^2 = \mu_X^{-2}(-\theta, 1)\begin{pmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix}\begin{pmatrix}-\theta\\ 1\end{pmatrix}$$

and an application of the delta method gives the result.

σ_θ^2 could be estimated by replacing μ_Y by $\bar{Y}$, σ_Y^2 by $n^{-1}\sum_j(Y_j - \bar{Y})^2$, $\rho\sigma_X\sigma_Y$ by $n^{-1}\sum_j(X_j - \bar{X})(Y_j - \bar{Y})$, etc.; note that as the variances are finite, all these expressions will converge in probability to the corresponding theoretical quantities.