PROBABILITY THEORY – FINAL EXAM SOLUTIONS

1. F, T, F, F, F, T, F, F, F, T.

2. (a) There are two possibilities:
   - one value appears 3 times, two values appear once ("three of a kind"):

$$\mathbb{P}(\text{Case 1}) = \frac{1}{6^5} \cdot \underbrace{\binom{6}{3}}_{\substack{\text{choosing}\\ \text{3 numbers}}} \cdot \underbrace{3}_{\substack{\text{choosing}\\ \text{the number}\\ \text{that appears 3x}}} \cdot \underbrace{\frac{5!}{3!}}_{\substack{\text{number of}\\ \text{permutations}\\ \text{of word abccc}}} = \frac{1200}{6^5};$$

   - one value appears once, two values appear twice ("two pairs"):

$$\mathbb{P}(\text{Case 2}) = \frac{1}{6^5} \cdot \underbrace{\binom{6}{3}}_{\substack{\text{choosing}\\ \text{3 numbers}}} \cdot \underbrace{3}_{\substack{\text{choosing}\\ \text{the number}\\ \text{that appears 1x}}} \cdot \underbrace{\frac{5!}{2!2!}}_{\substack{\text{number of}\\ \text{permutations}\\ \text{of word abbcc}}} = \frac{1800}{6^5}.$$

   So the total probability is $\frac{1200+1800}{6^5} = \frac{3000}{7776} \approx 38.58\%$.

   (b) The total number of arrangements is $n!$. The number of arrangements in which Alice is next to Bob and Charles is next to David is

$$\underbrace{(n-2)!}_{\substack{\text{treat A+B,}\\ \text{C+D as single}\\ \text{individuals}}} \cdot \underbrace{2}_{\text{order of A+B}} \cdot \underbrace{2}_{\text{order of C+D}} = 4(n-2)!$$

   The final answer is thus $\frac{4(n-2)!}{n!} = \frac{4}{n(n-1)}$.

   - We represent a placement of balls into urns as a permutation of the symbols:

$$\circ\circ\circ\circ\circ\circ\circ\circ\circ\circ\,||$$

   ($\circ$ represents a ball and $|$ represents a delimiter between two urns). The number of such permutations is $\frac{12!}{2!10!} = 66$.

3. (a)

$$F_Y(y) = \mathbb{P}(X \le \ln y) = F_X(\ln y) = \begin{cases} \ln(y) & \text{if } \ln(y) \in (0,1); \\ 0 & \text{if } \ln(y) \le 0; \\ 1 & \text{if } \ln(y) \ge 1. \end{cases} = \begin{cases} \ln(y) & \text{if } y \in (1,e); \\ 0 & \text{if } y \le 1; \\ 1 & \text{if } y \ge e. \end{cases}$$

$$\Longrightarrow f_Y(y) = \begin{cases} y^{-1} & \text{if } y \in (1,e); \\ 0 & \text{otherwise.} \end{cases}$$

   (b)

$$\mathbb{P}(X \ge x) = \sum_{i=x}^{\infty} f_X(i) \le \sum_{i=x}^{\infty} \frac{i}{xI} \cdot f_X(i) = \frac{1}{x} \sum_{i=x}^{\infty} i \cdot f_X(i) \le \frac{1}{x} \sum_{i=0}^{\infty} i \cdot f_X(i) = \frac{1}{x}\mathbb{E}(X).$$

(c) $M_U(t) = e^{\lambda_1(e^t-1)}$ and $M_V(t) = e^{\lambda_2(e^t-1)}$. Since they are independent,

$$M_{U+V}(t) = M_U(t) \cdot M_V(t) = e^{(\lambda_1+\lambda_2)(e^t-1)}.$$

This is the moment generating function of the Poisson distribution with parameter $\lambda_1 + \lambda_2$.

4. (a) For $i = 1, \ldots, 50$, let X_i be the indicator function of the event that man i recovers his own hat. Then,

$$X = \text{number of men who recover their own hats} = \sum_{i=1}^{50} X_i.$$

Each $X_i \sim \text{Bernoulli}(p)$ with

$$p = \mathbb{P}(X_i = 1) = \frac{1}{50},$$

so

$$\mathbb{E}(X_i) = \frac{1}{50}, \quad \text{Var}(X_i) = \frac{1}{50} \cdot \frac{49}{50} = \frac{49}{50^2}.$$

For $i \neq j$,

$$\mathbb{E}(X_i X_j) = \mathbb{P}(X_i = X_j = 1) = \frac{48!}{50!} = \frac{1}{50 \cdot 49},$$

so

$$\text{Cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) = \frac{1}{50 \cdot 49} - \frac{1}{50^2} = \frac{1}{50^2 \cdot 49}.$$

We then have

$$\mathbb{E}(X) = \sum_{i=1}^{50} \mathbb{E}(X_i) = 50 \cdot \frac{1}{50} = 1$$

and

$$\text{Var}(X) = \sum_{i=1}^{50} \text{Var}(X_i) + 2 \sum_{1 \le i < j \le 50} \text{Cov}(X_i, X_j) = 50 \cdot \frac{49}{50^2} + 2 \cdot \binom{50}{2} \cdot \frac{1}{50^2 \cdot 49} = 1.$$

(b) X is a sum of 50 Bernoulli random variables. They are not independent, but they are close to being independent: they have covariance close to zero (and recall that, for Bernoulli random variables, covariance equal to zero implies independence). Hence, the distribution of X is close to Binomial(50, 1/50). This is approximately equal to a Poisson distribution with parameter $\lambda = 50 \cdot \frac{1}{50} = 1$.

(c) $\mu_X = 1$, $\sigma_X = 1$, and by the Central Limit Theorem,

$$\sqrt{100} \cdot \frac{\bar{X}_{100} - \mu_X}{\sigma_X} \approx \mathcal{N}(0, 1).$$

Hence, letting $Z \sim \mathcal{N}(0, 1)$,

$$\begin{aligned} \mathbb{P}(\bar{X}_{100} > 1.08) = \mathbb{P}(10(\bar{X}_{100} - 1) > 0.8) &\approx \mathbb{P}(Z > 0.8) \\ &= 1 - \mathbb{P}(Z \le 0.8) \approx 1 - 0.7881 = 0.2119. \end{aligned}$$

5. (a)

$$C\int_0^1\int_0^2 (x+2y)\;\mathrm{d}x\;\mathrm{d}y = C\int_0^1 (2+4y)\;\mathrm{d}y = 4C \Longrightarrow C=\frac{1}{4}.$$

(b)

$$f_X(x)=\int_0^1 \frac{1}{4}(x+2y)\mathrm{d}y=\frac{1}{4}(x+1),\; 0<x<2.$$

(c) Let $(z,w)=g(x,y)=(x,x+2y)$, so that $h(z,w)=g^{-1}(z,w)=(z,(w-z)/2)$. We have $f_{Z,W}(z,w)>0$ if and only if $f_{X,Y}(h(z,w))>0$, that is, if $0<z<2$ and $z<w<z+2$. Also,

$$\left|\det\begin{pmatrix}\partial h_1/\partial z & \partial h_1/\partial w\\ \partial h_2/\partial z & \partial h_2/\partial w\end{pmatrix}\right| = \left|\begin{pmatrix}1 & 0\\ -1/2 & 1/2\end{pmatrix}\right| = \frac{1}{2}$$

and so

$$f_{X,W}(x,w)=f_{Z,W}(x,w)=f_{X,Y}(h(x,w))\cdot\frac{1}{2}=\frac{1}{8}\left(x+2\cdot\frac{w-x}{2}\right)=\frac{w}{8},$$

with $0<x<2$ and $x<w<x+2$.

6. (a)

$$\begin{aligned}\mathrm{Var}(XY)=\mathbb{E}(X^2Y^2)-\mathbb{E}(XY)^2 &= \mathbb{E}(X^2)\mathbb{E}(Y^2)-\mathbb{E}(X)^2\mathbb{E}(Y)^2\\ &=(\sigma_X^2+\mu_X^2)(\sigma_Y^2+\mu_Y^2)-\mu_X^2\mu_Y^2\\ &=\sigma_X^2\sigma_Y^2+\mu_Y^2\sigma_X^2+\mu_X^2\sigma_Y^2.\end{aligned}$$

(b)

$$\mathrm{Cov}(X_i-\bar{X}_n,\bar{X}_n)=\mathrm{Cov}(X_i,\bar{X}_n)-\mathrm{Cov}(\bar{X}_n,\bar{X}_n)=\mathrm{Cov}\left(X_i,\frac{1}{n}\sum_{j=1}^n X_j\right)-\mathrm{Var}(\bar{X}_n)$$

$$=\frac{1}{n}\sum_{j=1}^n \mathrm{Cov}(X_i,X_j)-\frac{\sigma^2}{n}=\frac{1}{n}\mathrm{Var}(X_i)-\frac{\sigma^2}{n}=0,$$

where we have used the fact that $\mathrm{Cov}(X_i,X_j)=0$ for $i\neq j$, due to independence.

7. (a)

$$\mathbb{E}(g(U_i))=\int_{-\infty}^{\infty} g(x)\cdot f_{U_i}(x)\;\mathrm{d}x=\int_0^1 g(x)\;\mathrm{d}x.$$

The random variables $X_1=g(U_1)$, $X_2=g(U_2)$, … are independent and identically distributed with mean $\int_0^1 g(x)\mathrm{d}x$. By the law of large numbers, $\frac{1}{n}\sum_{i=1}^x X_i$ converges in probability, as $n\to\infty$, to its mean $\int_0^1 g(x)\mathrm{d}x$.

(b) For $i=1,\ldots,1000$, let I_i be the indicator function of the event that rock i is inside the circle. These are independent Bernoulli random variables with

$$\mathbb{P}(I_i=1)=\iint_{\text{circle}} f_{X,Y}(x,y)\;\mathrm{d}x\mathrm{d}y.$$

Noting that

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{4} & \text{if } (x,y) \in \text{square} \\ 0 & \text{otherwise,} \end{cases}$$

we get

$$\mathbb{P}(I_i = 1) = \frac{\text{Area(circle)}}{4} = \frac{\pi}{4} = \mu_{I_i}.$$

By the law of large numbers, $\frac{1}{1000}\sum_{i=1}^{1000} I_i$ is close to $\mu_{I_i} = \frac{\pi}{4}$. Hence, we get an estimate for π by computing

$$4 \cdot \frac{1}{1000}\sum_{i=1}^{1000} I_i = \frac{\#\{\text{rocks inside the circle}\}}{250}.$$

8. $\log(Y) = W_1 + \cdots + W_n$, where

$$W_i = \log X_i = \begin{cases} \log(1) = 0 & \text{with probability } 1/10; \\ \log(10) = 1 & \text{with probability } \frac{9}{10} \end{cases}$$

(here, 'log' indicates the logarithm on base 10). The W_i's are then independent Bernoulli random variables with

$$\mu_{W_i} = \frac{9}{10}, \quad \sigma^2_{W_i} = \frac{9}{10} - \left(\frac{9}{10}\right)^2 = \frac{9}{100}.$$

Hence, the distribution of $\log(Y)$ is Binomial(400, 9/10), which can be approximated by $\mathcal{N}(360, 36)$. Letting $Z' \sim \mathcal{N}(360, 36)$ and $Z \sim \mathcal{N}(0,1)$, we have

$$\begin{aligned} \mathbb{P}\left(10^{354} \le Y \le 10^{363}\right) &= \mathbb{P}\left(354 \le \log(Y) \le 363\right) \\ &\approx \mathbb{P}\left(353.5 \le Z' \le 363.5\right) \\ &= \mathbb{P}\left(\frac{353.5 - 360}{6} \le Z \le \frac{363.5 - 360}{6}\right) \\ &= \mathbb{P}\left(-1.08 \le Z \le 0.58\right) \\ &= F_Z(0.58) - 1 + F_Z(1.08) \approx 0.7190 - 1 + 0.8599 = 0.5789. \end{aligned}$$