1. (WMS, Problem 3.118.) Five cards are dealt at random and without replacement from a standard deck of 52 cards. What is the probability that the hand contains all 4 aces if it is known that it contains at least 3 aces?

   Solution. Note that here the population consists of $N = 52$ cards, with $r = 4$ aces and the sample size is $n = 5$. Let X denote the number of aces in the hand. Then $X \sim \text{HG}(N = 52, r = 4, n = 5)$ which means that the PMF of X is

$$p(x) = \frac{\binom{4}{x}\binom{48}{5-x}}{\binom{52}{5}}; \quad x = 0, 1, 2, 3, 4.$$

   Hence, required probability:

$$P(X = 4 | X \geq 3) = \frac{P(\{X = 4\} \cap \{X \geq 3\})}{P(X \geq 3)} = \frac{P(X = 4)}{P(X \geq 3)} = \frac{\binom{4}{4}\binom{48}{1}}{\binom{4}{4}\binom{48}{1} + \binom{4}{3}\binom{48}{2}} = \boxed{0.0105.}$$

   $\square$

2. (WMS, Problem 3.135.) A salesperson has found that the probability of a sale on a single contact is approximately 0.03. If the salesperson contacts 100 prospects, what is the probability of making at least one sale?

   Solution. We can assume sales on different contacts to be independent. Let $X = \#$ sales. Then $X \sim \text{Bin}(n = 100, p = 0.03)$. Because $n = 100$ is large, $p = 0.03$ is small and $np = 3$ is moderate, we can approximate the $\text{Bin}(n = 100, p = 0.03)$ distribution by the $\text{Poi}(\lambda = np = 3)$ distribution. Therefore, required probability:

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) \approx 1 - e^{-3}\,\frac{3^0}{0!} = 1 - e^{-3} = \boxed{0.9502.}$$

   $\square$

3. An airline sells 200 tickets for a certain flight on an airplane that has only 198 seats because, on the average, 1 percent of purchasers of airline tickets do not appear for the departure of their flight. Determine the probability that everyone who appears for the departure of this flight will have a seat. [**Hint:** Define $X = \#$ people who do not appear for their flight.]

   Solution. As suggested in the hint, let $X = \#$ people who do not appear for their flight. We can assume that the passengers independently decide to show up or not, and we'll consider not appearing for the flight a success (for the airline!). Then X is the number of successes in a sequence of $n = 200$ independent Bernoulli trials with probability of success $p = 1\% = 0.01$. So, $X \sim \text{Bin}(n = 200, p = 0.01)$. Because $n = 200$ is large, $p = 0.01$ is small and $np = 2$ is moderate, we can approximate the $\text{Bin}(n = 200, p = 0.01)$ distribution by $\text{Poi}(\lambda = np = 2)$ distribution.

   Now, everyone will get a seat if and only if at least 2 passengers do not appear, i.e., $X \geq 2$. Therefore, required probability:

$$P(X \geq 2) = 1 - P(X \leq 1)$$

$$
\begin{aligned}
&= 1 - P(X=0) - P(X=1) \\
&\approx 1 - e^{-2}\frac{2^0}{0!} - e^{-2}\frac{2^1}{1!} \\
&= 1 - 3e^{-2} = \boxed{0.5940.}
\end{aligned}
$$

□

4. Suppose a discrete RV X with support $\mathscr{X} = \{-N, -(N-1), \cdots, -1, 0, 1, \cdots, N-1, N\}$ is symmetric (about 0), i.e., $P(X=k) = P(X=-k)$ for all k.
   (a) Show for any odd positive integer r, $\mu'_r = 0$, where μ'_r denotes the r-th raw moment of X.
   (b) Prove that for any k, $\mu_k = \mu'_k$, where μ_k denotes the k-th central moment of X. Thus, μ_r is also zero when r is an odd positive integer.

Solution. (a) Let $p(k) = P(X=k)$ for $k \in \mathbb{R}$. Then, $p(k) = p(-k)$ for all k. Therefore, for any odd positive integer r,

$$
\begin{aligned}
\mu'_r &= E(X^r) \\
&= \sum_{x \in \mathscr{X}} x^r p(x) = \sum_{x=-N}^{N} x^r p(x) = \sum_{x=-N}^{-1} x^r p(x) + 0 + \sum_{x=1}^{N} x^r p(x) \\
&= \sum_{y=1}^{N} \underbrace{(-y)^r}_{=-y^r} \underbrace{p(-y)}_{=p(y)} + \sum_{x=1}^{N} x^r p(x) && (y=-x) \\
&= -\sum_{x=1}^{N} x^r p(x) + \sum_{x=1}^{N} x^r p(x) = 0. && \text{QED.}
\end{aligned}
$$

(b) Since $r=1$ is an odd positive integer, therefore, from part (a), $\mu = E(X) = \mu'_1 = 0$. Hence $\mu_r = E(X-\mu)^r = E(X-0)^r = E(X^r) = \mu'_r$.

□

5. (WMS, Problem 3.147 - 3.148.) Let Y have a geometric distribution with probability of success p, and define $q = 1-p$.
   (a) Show that the MGF for Y is $M_Y(t) = \frac{pe^t}{1-qe^t}$.

   Solution. The MGF for Y is

$$
\begin{aligned}
M_Y(t) &= E(e^{tY}) \\
&= \sum_y e^{ty} p(y) = \sum_{y=1}^{\infty} e^{ty} q^{y-1} p \\
&= pe^t \sum_{y=1}^{\infty} e^{t(y-1)} q^{y-1} \\
&= pe^t \sum_{x=0}^{\infty} (qe^t)^x && x = y-1 \\
&= pe^t \frac{1}{1-qe^t}
\end{aligned}
$$

   provided $|qe^t| < 1 \iff e^t < 1/q \iff t < \log(1/q) = -\log q$. □

(b) Differentiate the MGF in part (a) to find $E(Y)$ and $E(Y^2)$. Then find $V(Y)$.

Solution. Verify that

$$M'_Y(t) = \frac{d}{dt}M_Y(t) = \frac{pe^t}{(1-qe^t)^2}$$
$$M''_Y(t) = \frac{d^2}{dt^2}M_Y(t) = \frac{pe^t(1+qe^t)}{(1-qe^t)^3}.$$

Therefore,

$$E(X) = M'_Y(t)\big|_{t=0} = \left.\frac{pe^t}{(1-qe^t)^2}\right|_{t=0} = \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}$$
$$E(X^2) = M''_Y(t)\big|_{t=0} = \left.\frac{pe^t(1+qe^t)}{(1-qe^t)^3}\right|_{t=0} = \frac{p(1+q)}{(1-q)^3} = \frac{p(1+q)}{p^3} = \frac{1+q}{p^2}.$$

Hence,

$$V(X) = E(X^2) - E^2(X) = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}.$$

□

6. First, a result: If X is a nonnegative RV with finite expectation and $a > 0$, then

$$P(X \geq a) \leq \frac{E(X)}{a} \quad \text{(Markov inequality).}$$

Using the above inequality, prove Chebyshev's theorem: if Y is a RV with mean μ and finite variance σ^2, then, for any constant $k > 0$,

$$P(|Y-\mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Solution. First, note that for a real number y, $|y-\mu| \geq k\sigma \iff (y-\mu)^2 \geq k^2\sigma^2$. Thus, the two sets $\{|Y-\mu| \geq k\sigma\}$ and $\{(Y-\mu)^2 \geq k^2\sigma^2\}$ are the same. Now if we let $X = (Y-\mu)^2$ then X is a nonnegative RV with expectation $E(X) = E(Y-\mu)^2 = V(Y) = \sigma^2 < \infty$. (We'll assume $\sigma > 0$.) Hence, we can apply Markov inequality on $X = (Y-\mu)^2$. Thus,

$$\begin{aligned} P(|Y-\mu| \geq k\sigma) &= P\left((Y-\mu)^2 \geq k^2\sigma^2\right) \\ &= P(X \geq k^2\sigma^2) && X = (Y-\mu)^2 \\ &\leq \frac{E(X)}{k^2\sigma^2} && \text{(Markov inequality)} \\ &= \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2} && \text{QED.} \end{aligned}$$

□

7. This exercise demonstrates the *tightness* of Chebyshev's theorem. For any constant k, define a RV X with support $\mathscr{X} = \{-1, 0, 1\}$ and PMF $p_X(-1) = p_x(1) = \frac{1}{2k^2}$ and $p_X(0) = 1 - \frac{1}{k^2}$.

   (a) Verify that p_X is indeed a PMF.

   (b) Find the mean μ_X and variance σ_X^2 of X.

(c) Show that $P(|X - \mu_X| \geq k\sigma_X) = \frac{1}{k^2}$. [**Note:** Thus, for any positive constant k, we can construct a RV X (and hence a probability distribution), for which equality holds in the Chebyshev theorem.]

Solution. (a) Clearly $p_X(x) \geq 0$ for all $x \in \mathbb{R}$, and

$$\sum_{x \in \mathscr{X}} p_X(x) = p_X(-1) + p_X(0) + p_X(1) = \frac{1}{2k^2} + 1 - \frac{1}{k^2} + \frac{1}{2k^2} = 1.$$

(b) Observe that X is symmetric (about 0). Therefore, from problem 4 in this homework, we get

$$\begin{aligned} \mu_X &= E(X) = 0 \\ \sigma_X^2 &= V(X) = E(X^2) = 1 \times \frac{1}{2k^2} + 0 + 1 \times \frac{1}{2k^2} = \frac{1}{k^2}. \end{aligned}$$

(c) From part (b), we have $\sigma_X = \frac{1}{k}$. Therefore,

$$\begin{aligned} P(|X - \mu_X| \geq k\sigma_X) &= P\left(|X| \geq k \cdot \frac{1}{k}\right) \\ &= P(|X| \geq 1) = P(X = \pm 1) \\ &= p(1) + p(-1) = \frac{1}{2k^2} + \frac{1}{2k^2} = \frac{1}{k^2}. \end{aligned}$$

□

8. A RV X with support $\mathscr{X} = \{1, 2, \cdots, N\}$ is said to follow a discrete uniform distribution (over the set of first N positive integers) if all elements of the support $\mathscr{X}$ are equally probable as values of X, i.e., if $P(X = x) = c$, for some constant c and for all $x \in \mathscr{X}$.

(a) Find c.

(b) Find $E(X)$ and $V(X)$. The following formulas should be helpful:

$$\begin{aligned} 1 + 2 + \cdots + N &= \frac{N(N+1)}{2} \\ 1^2 + 2^2 + \cdots + N^2 &= \frac{N(N+1)(2N+1)}{6} \end{aligned}$$

Solution. (a) From $\sum_{x \in \mathscr{X}} p(x) = 1$ we get, $\sum_{x=1}^{N} c = Nc = 1 \implies c = \frac{1}{N}$. Thus, the PMF of X is $p(x) = \frac{1}{N}$ for $x = 1, \cdots, N$.

(b) We have

$$E(X) = \sum_{x \in \mathscr{X}} xp(x) = \sum_{x=1}^{N} x\frac{1}{N} = \frac{1}{N} \cdot \frac{N(N+1)}{2} = \frac{N+1}{2}$$

and

$$E(X^2) = \sum_{x^2 \in \mathscr{X}} xp(x) = \sum_{x=1}^{N} x^2 \frac{1}{N} = \frac{1}{N} \cdot \frac{N(N+1)(2N+1)}{6} = \frac{(N+1)(2N+1)}{6}.$$

Hence,

$$\begin{aligned}
V(X) &= E(X^2) - E^2(X) \\
&= \frac{(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4} \\
&= \frac{2(N+1)(2N+1) - 3(N+1)^2}{12} \\
&= \frac{(N+1)(4N+2-3N-3)}{12} = \frac{(N+1)(N-1)}{12} = \frac{N^2-1}{12}.
\end{aligned}$$

□

9. (WMS, Problem 4.8.) Suppose that Y has PDF

$$f(y) = \begin{cases} ky(1-y), & 0 \leq y \leq 1 \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the value of k that makes $f(y)$ a probability density function.

(b) Find the CDF $F(y)$ of Y.

(c) Calculate $P(0.4 \leq Y < 1)$.

(d) Calculate $P(Y \leq 0.4 \,|\, Y \leq 0.8)$ and hence find $P(Y \geq 0.4 \,|\, Y \leq 0.8)$.

Solution. (a) Clearly $f(y) \geq 0$ for all y. Now, from $\int_{-\infty}^{\infty} f(y)\,dy = 1$ we get,

$$k\int_0^1 y(1-y)\,dy = k\left(\int_0^1 y\,dy - \int_0^1 y^2\,dy\right) = k\left(\frac{1}{2} - \frac{1}{3}\right) = \frac{k}{6} = 1 \implies \boxed{k = 6.}$$

(b) Clearly $F(y) = 0$ for $y < 0$ and $F(y) = 1$ for $y \geq 1$ as Y has support $\mathscr{Y} = [0,1]$. For $0 \leq y \leq 1$,

$$F(y) = \int_{-\infty}^{y} f(x)dx = 6\int_0^y x(1-x)dx = 6\left(\int_0^y x\,dx - \int_0^y x^2\,dx\right) = 6\left(\frac{y^2}{2} - \frac{y^3}{3}\right) = y^2(3-2y).$$

Thus, the DF of Y is given by:

$$F(y) = \begin{cases} 0, & y < 0 \\ y^2(3-2y), & 0 \leq y < 1 \\ 1, & y \geq 1. \end{cases}$$

(c) Note that,

$$\begin{aligned}
P(0.4 \leq Y < 1) &= P(0.4 < Y \leq 1) \qquad \text{(Y is continuous)} \\
&= F(1) - F(0.4) \\
&= 1 - (0.4)^2(3 - 2 \times 0.4) = 1 - 0.352 = \boxed{0.648.}
\end{aligned}$$

(d) Note that,

$$P(Y \leq 0.4 \,|\, Y \leq 0.8) = \frac{P(\{Y \leq 0.4\} \cap \{Y \leq 0.4\})}{P(Y \leq 0.8)}$$

$$= \frac{P(Y \le 0.4)}{P(Y \le 0.8)} = \frac{(0.4)^2(3 - 2 \times 0.4)}{(0.8)^2(3 - 2 \times 0.8)} = \frac{0.352}{0.896} \approx \boxed{0.3928.}$$

Also,

$$\begin{aligned} P(Y \ge 0.4 \mid Y \le 0.8) &= P(Y > 0.4 \mid Y \le 0.8) \qquad \text{(Y is continuous)} \\ &= 1 - P(Y \le 0.4 \mid Y \le 0.8) \approx 1 - 0.3928 = \boxed{0.6072.} \end{aligned}$$

□

10. (WMS, Problem 4.19.) Let the DF of a random variable Y be

$$F(y) = \begin{cases} 0, & y \le 0 \\ \frac{y}{8}, & 0 < y < 2 \\ \frac{y^2}{16}, & 2 \le y < 4 \\ 1, & y \ge 4. \end{cases}$$

(a) Find the PDF of Y.

(b) Find $P(1 \le Y \le 3)$.

(c) Find $P(Y \ge 1.5)$.

(d) Find $P(Y \ge 1 | Y \le 3)$.

Solution. (a) Observe that Y has support $\mathscr{Y} = (0, 4)$. So, if f denotes the PDF of Y, then $f(y) = 0$ for $y \notin \mathscr{Y} = (0, 4)$. Thus Y has PDF,

$$f(y) = \begin{cases} \frac{d}{dy}\frac{y}{8} = \frac{1}{8}, & 0 < y < 2 \\ \frac{d}{dy}\frac{y^2}{16} = \frac{2y}{16} = \frac{y}{8}, & 2 \le y < 4 \\ 0, & \text{otherwise.} \end{cases}$$

(b) We have,

$$\begin{aligned} P(1 \le Y \le 3) &= P(1 < Y \le 3) \qquad \text{(Y is continuous)} \\ &= F(3) - F(1) \\ &= \frac{3^2}{16} - \frac{1}{8} = \frac{9}{16} - \frac{2}{16} = \boxed{\frac{7}{16}.} \end{aligned}$$

(c) Note that

$$P(Y \ge 1.5) = 1 - P(Y < 1.5) = 1 - P(Y \le 1.5) = 1 - F(1.5) = 1 - \frac{1.5}{8} = 1 - \frac{3}{16} = \boxed{\frac{13}{16}.}$$

(d) From part (a), $P(1 \le Y \le 3) = 7/16$. Therefore,

$$P(Y \ge 1 | Y \le 3) = \frac{P(1 \le Y \le 3)}{P(Y \le 3)} = \frac{7/16}{9/16} = \boxed{\frac{7}{9}.}$$

□