Problem 1

Solution

Let X_i be a random variable. If the i-th bin is not empty, then $X_i = 1$. Ohterwise, $X_i = 0$. For all i, we have $P\{X_i = 1\} = 1 - (1 - \frac{1}{n})^m$. Thus, the expected number of non-empty bins is $\mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}\left[X_i\right] = n(1 - (1 - \frac{1}{n})^m)$.

Problem 2

Solution

- Let $X \sim Ber(0.5)$. PMF of X is $P\{X = 0\} = 0.5; P\{X = 1\} = 0.5$. MGF of X is $M(\theta) = \mathbb{E}\left[e^{\theta X}\right] = 0.5 + 0.5e^{\theta}$.
- Let $X \in \mathbb{N}^+$ be a random variable, and $P\{X = i\} = \frac{1}{2^i},\ i = 1, 2, 3, \dots$. Then, $M(\theta) = \sum_{i=1}^{\infty}(\frac{e^{\theta}}{2})^i$, which only exists when $\theta \in (-\infty, \ln 2)$.
- Let $X \in \{\pm 2^i,\ i = 1, 2, 3, \dots\}$ be a random variable, and $P\{X = 2^i\} = \frac{1}{2^{i+1}},\ P\{X = -2^i\} = \frac{1}{2^{i+1}},\ i = 1, 2, 3, \dots$. Then, $M(\theta) = \sum_{i=1}^{\infty} \frac{1}{2^{i+1}}(e^{2^i\theta} + e^{-2^i\theta}) > \sum_{i=1}^{\infty} \frac{e^{2^i|\theta|}}{2^{i+1}}$. The term in the summation $\frac{e^{2^i|\theta|}}{2^{i+1}}$ diverges to ∞ when $\theta \neq 0$. Thus, $M(\theta)$ only exists when $\theta = 0$.
- (i) Yes. Let X, Y be i.i.d. random variables. If $M(\theta)$ is the MGF of X, then the MGF of $X+Y$ is $M(\theta)^2$. (ii) Yes. $M(k\theta) = \mathbb{E}\left[e^{k\theta X}\right]$ which is the MGF of kX. (iii) No. $kM(0) = k \neq 1$. (iv) Yes. $e^{k\theta}M(\theta) = \mathbb{E}\left[e^{\theta(X+k)}\right]$ which is the MGF of $X + k$.

Problem 3

Solution

- Markov's inequality: If Y is a nonnegative random variable, then for $c > 0$,
$$P\{Y \geq c\} \leq \frac{\mathbb{E}\left[Y\right]}{c}.$$
- For $\theta \geq 0$,
$$\begin{aligned} P\{\sum_{i=1}^{n} X_i \geq nx\} &= P\{e^{\theta(\sum_{i=1}^{n} X_i - nx)} \geq 1\} \\ &\leq \mathbb{E}\left[e^{\theta(\sum_{i=1}^{n} X_i - nx)}\right] = \mathbb{E}\left[e^{\theta X_1}\right]^n e^{-\theta nx} = e^{-n(\theta x - \ln \mathbb{E}\left[e^{\theta X_1}\right])} \leq e^{-n \sup_{\theta > 0}(\theta x - \ln \mathbb{E}\left[e^{\theta X_1}\right])} \end{aligned}$$
- This only holds for $x > p$. Let $g(\theta) = \theta x - \log(\mathbb{E}\left[e^{\theta X}\right]) = \theta x - \log(1 - p + pe^{\theta})$. Sovle $g'(\theta) = 0$, and we have $\theta^* = \log(\frac{x}{p} \cdot \frac{1-p}{1-x})$. Because $x > p$, $\theta^* > 0$, and $\sup_{\theta > 0} g(\theta) = g(\theta^*) = \mathbb{D}(x||p)$.

Problem 4

Solution

- *Positive definite matrix:* A symmetric real-valued matrix A is said to be positive definite if and only if for all column vector $x \neq \vartheta$, $x^\intercal Ax > 0$. Such a matrix has only positive eigenvalues. For such a matrix, a decomposition $A = Q^\intercal \Lambda Q$ is always possible, where $Q^\intercal Q = I$ and Λ is a diagonal matrix with only positive elements. Thus, we can decompose Σ by $\Sigma = Q^\intercal \Lambda Q$. Let $A = \sqrt{\Lambda}^{-1}Q$, where $\sqrt{\Lambda}$ is the element-wise square root of Λ, and $\sqrt{\Lambda}^{-1}$ is valid because Σ has only positive eigenvalues. Because Y is a finite linear combination of X, Y is also a Gaussian random vector. Then, we have $\mathbb{E}[Y] = A(\mathbb{E}[X] - \mu) = 0$, and

$$\begin{aligned} Cov(Y) = \mathbb{E}[YY^\intercal] = \mathbb{E}[A(X-\mu)(X-\mu)^\intercal A^\intercal] &= A\Sigma A^\intercal \\ &= \sqrt{\Lambda}^{-1} Q\Sigma Q^\intercal \sqrt{\Lambda}^{-1} = \sqrt{\Lambda}^{-1}\Lambda\sqrt{\Lambda}^{-1} = I. \end{aligned}$$

  Thus, $Y \sim N(0, I)$.

- $A = \begin{bmatrix} 1 & 10 & 0 \\ 10 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ cannot be a covariance matrix. The 2nd principal minor of A is negative, so A is not positive definite. On the other hand, the first two random variables do not satisfy the Cauchy–Schwarz inequality, i.e. $Var(X_1)Var(X_2) < Cov(X_1, X_2)^2$.

Problem 5

Solution

- Jensen's inequality: Let ϕ be a convex function and let X be a random variable such that $\mathbb{E}[X]$ is finite. Then, $\mathbb{E}[\phi(X)] \geq \phi(\mathbb{E}[X])$.

- Let $\phi(X) = -\log(X), X > 0$. Then ϕ is a convex function. Let X be a random variable, and $P\{X = \frac{q_i}{p_i}\} = p_i$. Then, $\mathbb{D}(p||q) = \mathbb{E}[\phi(X)] \geq \phi(\mathbb{E}[X]) = \phi(\sum_{i=1}^m p_i \frac{q_i}{p_i}) = \phi(1) = 0$.