1. An urn contains 600 red balls and 400 black balls. A random sample of size 50 is taken without replacement. Let \( X \) denote the number of red balls in the sample. Show that the probability mass function of \( X \) is unimodal.

2. Hidden inside each box of Primate brand breakfast cereal is a small plastic figure of an animal: an ape, a baboon, or a chimp. Suppose a fraction \( \alpha \) of the very large population of cereal boxes contain apes, a fraction \( \beta \) contain baboons, and a fraction \( \gamma \) contain chimps. Find the expected number of boxes you need to buy before you have at least one figure of each type.

3. Let \( \{y_{ij} : i = 1, \ldots, r, j = 1, \ldots, c\} \) be a set of observations with the unusual property that \( y_{ij} = 1 \) if \( i = j = 1 \) and 0 otherwise. Suppose we fit an additive model to the data, by finding values \( \hat{t}, \hat{a}_i, \hat{b}_j \) to minimize \( \sum_{ij} (y_{ij} - t - a_i - b_j)^2 \).

   (a) Find the values of \( \hat{t}, \) the \( \hat{a}_i \)'s and the \( \hat{b}_j \)'s if they are subjected to the “sum constraints” \( \sum_i \hat{a}_i = 0 = \sum_j \hat{b}_j \).

   (b) Find the values of \( \hat{t}, \) the \( \hat{a}_i \)'s and the \( \hat{b}_j \)'s if they are subjected to the “treatment constraints” \( \hat{a}_1 = 0 = \hat{b}_1 \).

4. Three random points are chosen independently from the uniform distribution on a disc of unit radius. Find the probability that the center of the disc lies in the convex hull of the three points.

5. Suppose \( X_0, X_1, \ldots, X_n \) is a Markov chain on a finite state space, with transition matrix \( P \) and initial distribution \( \mu \). Define \( Y_i = X_{n-i} \) for \( i = 0, 1, \ldots, n \).

   (a) Show that \( \{Y_i : i = 0, 1, \ldots, n\} \) is also a Markov chain, with transition probabilities that might not be stationary.

   (b) If \( \mu \) is the stationary distribution for \( P \), show that the \( Y \)-chain also has stationary transition probabilities.

6. Consider the choice of a set estimator \( C_X \) for a parameter \( \theta \in \mathbb{R} \) based on one observation \( X \) from the \( N(\theta, 1) \) distribution, using the loss function

\[
L(\theta, C) = (\text{Lebesgue measure of } C) - 1\{\theta \in C\}
\]

Show that the set \( C_X = [X - c_0, X + c_0] \) is minimax for a suitably chosen constant \( c_0 \).
7. Suppose $T$ is an unobserved random variable with density $g(t) = \frac{1}{2} t^2 e^{-t} \{t > 0\}$, which is generated independently of a random variable $B$ for which $\mathbb{P}\{B = +1\} = 1/2 = \mathbb{P}\{B = -1\}$. We observe $B$ and $X = \theta + BT$ for an unknown $\theta \in \mathbb{R}$.

(a) Find the Fisher information function for a single observation $(X, B)$.

(b) Suppose we only observe $X$ and not $B$. Show that the Fisher information is the same as for part (a).

(c) If $(X, B)$ were observed, would $X$ be a sufficient statistic for $\theta$?

8. Let $\mathbb{P}_\theta$ denote the uniform distribution on $[0, \theta]^2$, for $\theta > 0$. That is, the coordinates $x_1$ and $x_2$ are independent Uniform$[0, \theta]$ under $\mathbb{P}_\theta$. Let $S := x_1 + x_2$ and $M := \max(x_1, x_2)$. Consider estimation of $\theta$ with loss function $L(\theta, a) := (\theta - a)^2$.

(a) Explain why $E_{\theta}(S \mid M = m)$ is preferred to $S$ for estimating $\theta$.

(b) Explain why $E_{\theta}(2x_1 \mid S)$ is preferred to $2x_1$ for estimating $\theta$.

(c) Explain why $E_{\theta}(3M/2 \mid S = s)$ is not preferred to $3M/2$ for estimating $\theta$.

9. Suppose $X_n$ has a Bin$(n, p_n)$ distribution with variance $\sigma_n^2 = np_n(1 - p_n)$ that converges to 1 as $n$ tends to infinity. Show that $(X_n - np_n)/\sigma_n$ cannot converge in distribution to $\mathcal{N}(0, 1)$.

10. Suppose $Z_1, \ldots, Z_k$ are independent random vectors, each distributed $\mathcal{N}(0, I_n)$. Let $u_0$ be a fixed unit vector in $\mathbb{R}^n$. Show that the squared length of the component of $u_0$ in the subspace spanned by $Z_1, \ldots, Z_k$ is distributed like $A/(A + B)$ where $A$ and $B$ are independent with $A \sim \chi^2_k$ and $B \sim \chi^2_{n-k}$.

11. Let $P$ and $Q$ be probability measures on a set $\mathcal{X}$ and $f$ be a measurable function from a set $\mathcal{X}$ into another set $\mathcal{Y}$. Write $\tilde{P}$ for the distribution of $f$ under $P$ and $\tilde{Q}$ for its distribution under $Q$. [You may assume $\mathcal{X}$ and $\mathcal{Y}$ are finite if you wish to avoid measure theoretic details.]

(a) Show that the relative entropy $D(\tilde{P} \mid \mid \tilde{Q})$ is less than or equal to $D(P \mid \mid Q)$.

(b) Let $\tilde{P}$ denote the Bin$(n, p)$ distribution and $\tilde{Q}$ denote the Poisson$(np)$ distribution. Show that $D(\tilde{P} \mid \mid \tilde{Q}) = O(np^2)$.

12. Suppose an urn initially contains $r$ red balls and $b$ black balls. At step $n$ a ball is selected at random from the urn, then replaced by $d_n$ balls of the same color, where $d_n$ is a positive random integer that might depend on the outcomes of the first $n-1$ draws. After completion of the $n$th step, let $R_n$ denote the number of red balls and $B_n$ the number of black balls in the urn. Show that $X_n := R_n/(R_n + B_n)$ is a martingale with respect to a suitable filtration.