

Part I

Define the Δ operation (symmetric difference) between two events (sets) E and F as:

$$E\Delta F = (E \cap F^c) \cup (E^c \cap F).$$

So, $E\Delta F$ is the set of outcomes (elements) that are in either E or F , but not in both.

1. Prove: $|P(A) - P(B)| \leq P(A\Delta B)$.
2. Three sets A, B , and C satisfy $A\Delta C = B\Delta C$. Show that $A = B$.
3. Recall that $d(x, y)$ is a metric if it satisfies: (i) $d(x, y) \geq 0$ and $d(x, y) = 0$ if and only if $x = y$, (ii) $d(x, y) = d(y, x)$, and (iii) $d(x, z) \leq d(x, y) + d(y, z)$. Assuming $P(A) = 0$ implies $A = \emptyset$ (empty set), argue that $d(A, B) := P(A\Delta B)$ is a metric.

Part II

Let X_1, X_2, \dots be a sequence of iid random variables taking only positive integer values, and $E(X_1) = a$. Define $S_n = \sum_{k=1}^n X_k$, for $n = 1, 2, \dots$

4. Show $E\left(\frac{S_m}{S_n}\right) = \frac{m}{n}$ for $1 \leq m \leq n$.
5. Show $E\left(\frac{1}{S_n}\right) < \infty$ and $E\left(\frac{S_m}{S_n}\right) = 1 + (m-n)aE\left(\frac{1}{S_n}\right)$ for $1 < n < m$.
6. Show $E\left(\frac{S_m}{S_n}\right) \geq \frac{m}{n}$ for all $m, n \geq 1$. (Prove and use the fact that $x + \frac{1}{x} \geq 2$, for $x > 0$.)

Part III

Let X and Y have the joint pdf

$$f(x, y) = \exp\left(c + 4x + 4y - \frac{x^2}{2} - \frac{y^2}{2} - \frac{x^2y^2}{2}\right)$$

for $-\infty < x, y < \infty$, for a constant c .

7. Find the marginal pdf of X .
8. Find the conditional pdf of X given $Y = y$. Identify the conditional distribution.

Part IV

Let X_1, X_2, \dots be a sequence of iid random variables with mean μ and variance $\sigma^2 < \infty$.

9. Argue that $\sqrt{n}(\bar{X}^2 - \mu^2) \xrightarrow{d} N(0, 4\mu^2\sigma^2)$, as $n \rightarrow \infty$.
10. Comment on the limit in Problem 9 when $\mu = 0$.
11. If $\mu = 0$, find a scaling factor a_n such that $a_n\bar{X}^2$ converges in distribution to a non-trivial limit.

Part 1

1. First note that $A \subset (A\Delta B) \cup B$. Using properties of probability, this gives $P(A) \leq P(A\Delta B) + P(B)$. Analogously, we get $P(B) \leq P(A\Delta B) + P(A)$. Rearranging terms we get $|P(A) - P(B)| \leq P(A\Delta B)$.
2. We will prove $A \subset B$ here (other case can be proved similarly). Let $x \in A$, then -

If $x \in C$: Then $x \in A \cap C$ and hence $x \notin A\Delta C = B\Delta C$. So, $x \in C$ but $x \notin B\Delta C$. Hence $x \in B \cap C \subset B$. This proves $A \subset B$.

If $x \in C^c$: Then $x \in A \cap C^c \subset A\Delta C = B\Delta C$. So, $x \notin C$ but $x \in B\Delta C = (B \cap C^c) \cup (C \cap B^c)$. That means $x \in B \cap C^c \subset B$. This proves $A \subset B$.
3. Using the symmetry in definition of the symmetric difference operation, we get that $d(A, B) = P(A\Delta B) = P(B\Delta A) = d(B, A)$. Also $d(A\Delta A) = P(\emptyset) = 0$ and $P(A\Delta B) = 0$ implies $A\Delta B = \emptyset$ (assuming $P(A) = 0$ implies $A = \emptyset$) and hence $A = B$. we will now prove: Prove: $|P(A) - P(B)| \leq P(A\Delta B)$, which will give $d(A, C) = P(A\Delta C) \leq P(A\Delta B) + P(B\Delta C) = d(A, B) + d(B, C)$, and complete the proof.

To show: Prove: $P(A\Delta C) \leq P(A\Delta B) + P(B\Delta C)$:

Choose $x \in A\Delta C = (A \cap C^c) \cup (A^c \cap C)$. We will deal with the case when $x \in (A \cap C^c)$ and show that x belongs to the right side (similar steps will prove the case when $x \in (A^c \cap C)$). So we have x in A but not in C :

If $x \in B$: then $x \in B \cap C^c \subset B\Delta C \subset (A\Delta B) \cup (B\Delta C)$.

If $x \in B^c$: then $x \in A \cap B^c \subset A\Delta B \subset (A\Delta B) \cup (B\Delta C)$.

Part 2

4. Note that $\frac{X_1}{S_n}, \dots, \frac{X_n}{S_n}$ are identically distributed (not independent) random variables. So, $\sum_{i=1}^n E\left(\frac{X_i}{S_n}\right) = nE\left(\frac{X_i}{S_n}\right)$ for $i = 1, \dots, n$. Also, using linearity of the expectation operation, we have: $\sum_{i=1}^n E\left(\frac{X_i}{S_n}\right) = E\left(\frac{S_n}{S_n}\right) = 1$. Hence, for $i = 1, \dots, n$, we have

$$nE\left(\frac{X_i}{S_n}\right) = 1 \implies E\left(\frac{X_i}{S_n}\right) = \frac{1}{n}$$

Hence for $1 \leq m \leq n$, $E\left(\frac{S_m}{S_n}\right) = \sum_{i=1}^m E\left(\frac{X_i}{S_n}\right) = \frac{m}{n}$.

5. Note that $X_i \geq 1$ for all $i = 1, 2, \dots$, So $S_n \geq 1$ as well as $\frac{1}{S_n} \leq 1$ and hence the expectation exists. For $1 < n < m$,

$$E\left(\frac{S_m}{S_n}\right) = \frac{S_n + \sum_{i=n+1}^m X_i}{S_n} = 1 + (m-n)E(X_i)E\left(\frac{1}{S_n}\right) = 1 + (m-n)aE\left(\frac{1}{S_n}\right)$$

using the fact that S_n and X_{n+1}, \dots, X_m are independent and that $E(X_i) = a$.

6. First note that $(x - 1)^2 \geq 0$ for all x . Expanding this and dividing by $x > 0$ yields, $x + 1/x \geq 2, x > 0$. In Problem 4 above, we showed for $1 \leq m \leq n$, $E\left(\frac{S_m}{S_n}\right) = \frac{m}{n}$. It remains to verify that for $n < m$, $E\left(\frac{S_m}{S_n}\right) \geq \frac{m}{n}$. First note that using the fact that $f(x) = 1/x, x > 0$ is convex, Jensen's inequality gives $E\left(\frac{1}{S_n}\right) \geq \frac{1}{E(S_n)} = \frac{1}{na}$. This and the statement proved in Problem 5 above, we have for $n < m$,

$$E\left(\frac{S_m}{S_n}\right) = 1 + (m-n)aE\left(\frac{1}{S_n}\right) \geq 1 + (m-n)a\frac{1}{na} = 1 + \frac{(m-n)}{n} = \frac{m}{n}.$$

Part 3

7. We will show the marginal $f_Y(y)$ by integrating over x (by symmetry, $f_X(x)$ has the same form). Fix $-\infty < y < \infty$,

$$\begin{aligned} f(x, y) &= \int_{-\infty}^{\infty} \exp\left(c + 4x + 4y - \frac{x^2}{2} - \frac{y^2}{2} - \frac{x^2y^2}{2}\right) dx \\ &= \exp\left(c + 4y - \frac{y^2}{2}\right) \int_{-\infty}^{\infty} \exp\left(4x - \frac{(1+y^2)x^2}{2}\right) dx \\ &= \exp\left(c + 4y - \frac{y^2}{2} + \frac{8}{1+y^2}\right) \int_{-\infty}^{\infty} \exp\left\{-\frac{(1+y^2)}{2} \left(x - \frac{4}{(1+y^2)}\right)^2\right\} dx \\ &= \frac{\sqrt{2\pi}}{1+y^2} \exp\left(c + 4y - \frac{y^2}{2} + \frac{8}{1+y^2}\right), \end{aligned}$$

using the fact that the rest of the terms integrate to 1 (total probability for a suitable Gaussian density function).

8. Dividing the joint density by the suitable marginal, we get for each $-\infty < x, y < \infty$ conditional pdf of X given $Y = y$ is

$$\begin{aligned} f(x|y) = \frac{f(x,y)}{f_Y(y)} &= \frac{\sqrt{1+y^2}}{\sqrt{2\pi}} \exp\left(4x - \frac{x^2}{2} - \frac{x^2y^2}{2} - \frac{8}{1+y^2}\right) \\ &= \frac{\sqrt{1+y^2}}{\sqrt{2\pi}} \exp\left\{-\frac{1+y^2}{2} \left(x - \frac{4}{1+y^2}\right)^2\right\}, \end{aligned}$$

by completing squares. So X given $Y = y$ is normal with mean $4/(1+y^2)$ and variance $1/(1+y^2)$ (even if the joint pdf was not bivariate normal).

Part 4

9. Using central limit theorem (CLT) for $\{X_i, \dots\}$, we get $\sqrt{n}(\bar{X} - \mu) \xrightarrow{d} N(0, \sigma^2)$. Using the delta method with $g(x) = x^2$ ($g'(x) = 2x$), we get that $\sqrt{n}(\bar{X}^2 - \mu^2) \xrightarrow{d} N(0, 4\mu^2\sigma^2)$, as $n \rightarrow \infty$.
10. When $\mu = 0$, Problem 9 gives $\sqrt{n}(\bar{X}^2 - 0)$ converges in law to the degenerate distribution at 0, as $n \rightarrow \infty$. And since, the limit is a constant, this convergence holds in probability as well.
11. The answer above suggests that we need a higher scaling factor a_n such that $a_n(\bar{X}^2)$ converges in distribution to a non-trivial asymptotic distribution. To find such a_n and the limit distribution, note that in this case ($\mu = 0$), CLT gives $\sqrt{n}(\bar{X} - 0) \xrightarrow{d} N(0, \sigma^2)$. Using continuous mapping theorem, this gives $n(\bar{X}^2) \xrightarrow{d} \sigma^2\chi_1^2$ as $n \rightarrow \infty$. So $a_n = n$ and limit distribution is $\sigma^2\chi_1^2$.

You may use the following facts on this Basic Theory II question set.

A random variable W with a hypergeometric distribution has probability mass function (pmf) given by

$$P(W = w) = \frac{\binom{M_1}{w} \binom{M_2}{m-w}}{\binom{M_1+M_2}{m}}, \quad \text{for integer } w \text{ such that } \max\{0, m - M_2\} \leq w \leq \{M_1, m\},$$

where M_1, M_2, m are non-negative integers with $1 \leq m \leq M_1 + M_2$.

A random variable Y with the ABC(a, b, c) distribution has pmf given by

$$P(Y = y) = \frac{\binom{a+y}{a} \binom{b+c-y}{b}}{\binom{a+b+c+1}{a+b+1}}, \quad y = 0, \dots, c,$$

based on non-negative integer parameters $a, b, c \geq 0$.

Part I (*Be sure to justify your answers, stating any standard results that you use.*)

Consider an urn with $N > 1$ balls, of which k balls are red while the remaining $N - k$ balls are blue. Suppose that N is *known*, but k is *unknown*. For inference about k , a random sample of n balls is drawn without replacement from the urn. Let X denote the number of red balls in the sample. The sample size n is assumed to be *known* with $1 \leq n \leq N$.

1. Describe the parameter space for k , denoted as (say) Θ_k .
2. Based on a given outcome $X = x \in \{0, \dots, n\}$, prove that the likelihood function $L(k)$ for $k \in \Theta_k$ satisfies $L(k) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$ when $x \leq k \leq N - (n - x)$ and $L(k) = 0$ otherwise.
3. For k with $x \leq k < N - (n - x)$, based again on given outcome $X = x \in \{0, \dots, n\}$, show that

$$\frac{L(k)}{L(k+1)} = \frac{(N-k)(k+1-x)}{(N-k-n+x)(k+1)}.$$

Hint: Use $\binom{b+1}{a} = \binom{b}{a} \frac{b+1}{b+1-a}$ for non-negative integer $a \leq b$.

4. Using **Problem 3**, show the following relation holds for k with $x \leq k < N - (n - x)$:

$$\frac{L(k)}{L(k+1)} > 1 \quad \text{if and only if} \quad k > k^* \equiv \frac{(N+1)x}{n} - 1.$$

Note: For reference in subsequent **Problems 5–6**, you may use that analogous relations also hold upon replacing each instance of “ $>$ ” with either “ $<$ ” or “ $=$ ” above.

5. If $k^* < x$ holds for $k^* \equiv \frac{(N+1)x}{n} - 1$ from **Problem 4** and an outcome $x \in \{0, \dots, n\}$ of X , prove that the likelihood $L(k)$ is a strictly decreasing, positive function of integer $k \in [x, N - (n - x)]$.

6. Using **Problems 3–5** to justify your answers, determine all potential values for the maximum likelihood estimator \hat{k}_{MLE} of k in each case **(a)** and **(b)** below based on an outcome $x \in \{0, \dots, n\}$ of X :

a) $x = 0$.

b) $x \leq k^* < N - (n - x)$ holds where $k^* \equiv \frac{(N+1)x}{n} - 1$ is an integer.

7. Considering a prior distribution on k that is $ABC(a, b, N)$ for some $a, b \geq 0$, where the form of the ABC distribution is defined on **Page 1**, show that the posterior distribution of $k - x$ given $X = x \in \{0, \dots, n\}$ is $ABC(a + x, b + (n - x), N - n)$, that is,

$$P(k - x = y|x) = \frac{\binom{a+x+y}{a+x} \binom{b+N-x-y}{b+(n-x)}}{\binom{a+b+N+1}{a+b+n+1}}, \quad y = 0, \dots, N - n.$$

Hint: Define the posterior probabilities $P(k = z|x)$ of k directly (over support outcomes $z \in \{x, \dots, x + N - n\}$) and show these are proportional to what is obtained by substituting “ $x + y$ ” with “ z ” in the form $P(k = y + x|x) = P(k = z|x)$ above for the ABC distribution.

8. Considering a pair of random variables (T, p) such that, for given $x \in \{0, \dots, n\}$,

the conditional distribution of T given p is $\text{Binomial}(N - n, p)$ with pmf

$$P(T = t|p) = \binom{N - n}{t} p^t (1 - p)^{N - n - t}, \quad t = 0, \dots, N - n,$$

and the marginal distribution of p is $\text{Beta}(x + 1, n - x + 1)$ with density function

$$f(p) = \frac{[x + (n - x)]!}{x!(n - x)!} p^x (1 - p)^{n - x}, \quad 0 < p < 1, \quad \text{and mean } \frac{x + 1}{(x + 1) + (n - x + 1)},$$

show that the marginal distribution of T matches the posterior distribution of $k - x$ from **Problem 7** when $a = b = 0$ there.

9. Referring to **Problems 7–8** as needed, show that under squared error loss and a prior for k that is $ABC(0, 0, N)$ the Bayes estimator of k is given by

$$\frac{(N + 2)(X + 1)}{n + 2} - 1.$$

Hint: Given x , the distribution of $T + x$ in **Problem 8** is the same as the posterior distribution of k , and moments of T can be found by conditioning on p in **Problem 8**.

Part II (*Be sure to justify your answers, stating any standard results that you use.*)

Consider the same urn from **Part I** containing N balls, of which k balls are red and $N - k$ balls are blue. Now let X denote the number of red balls found in a random sample of size n drawn *with replacement* from the urn; that is, $X = \sum_{i=1}^n X_i$ holds based on iid random variables X_1, \dots, X_n , where $X_i \in \{0, 1\}$ denotes the number of red balls found in the i th single random draw from the urn, $i = 1, \dots, n$. Note, however, that X is the *available observation* here, while X_1, \dots, X_n are not reported/available.

Again assume that the total number $N > 1$ of balls in the urn as well as the sample size $1 < n \leq N$ are *known*, while the number k of red balls is the *unknown* parameter of interest.

- 10.** Show that the method of moments (MOM) estimator, as given by $\hat{k}_{\text{MOM}} = NX/n$, is sufficient for k .

- 11.** Prove that the statistic defined below as a function of the MOM estimator \hat{k}_{MOM} , is unbiased for the parametric function $k(N - k)$.

$$\frac{n}{(n-1)} \cdot \hat{k}_{\text{MOM}}(N - \hat{k}_{\text{MOM}}).$$

Hint: In defining the observable X , the underlying X_1, \dots, X_n are iid Bernoulli(k/N) variables.

- 12.** If $L(k)$ denotes the likelihood function for k based on X , show that the ratio $L(k_1)/L(0)$ is a non-decreasing function of X when evaluated at a fixed $k_1 > 0$ in the parameter space of k .

- 13.** Using **Problem 12** or otherwise, for a given $\alpha \in (0, 1)$, carefully derive the uniformly most powerful (UMP) test of size α for evaluating $H_0 : k = 0$ vs. $H_1 : k > 0$.

- 14.** For a given size $\alpha \in (0, 1)$, show that the type II error level of the test from **Problem 13** is bounded above by $(1 - \alpha)(1 - N^{-1})^n$ under any alternative parameter value $k > 0$.

- 15.** Argue carefully that, under any alternative parameter value k with $k < N$, it holds that

$$\frac{\sqrt{n}(\hat{k}_{\text{MOM}} - k)}{\sqrt{N - \hat{k}_{\text{MOM}}}} \xrightarrow{d} \text{Normal}(0, k), \quad \text{as } n \rightarrow \infty.$$

That is, convergence in distribution holds to a normal limit with mean zero and variance k .

- 16.** By approximating $\sqrt{n}(\hat{k}_{\text{MOM}} - k)$ with a normal distribution having mean zero and variance $k(N - \hat{k}_{\text{MOM}})$ based on **Problem 15**, determine an approximate 95% confidence interval for k having a symmetric form $A_n \pm C_n$ for some statistics A_n, C_n .

Part I

1. The parameter space $\Theta_k = \{0, \dots, N\}$ for k is given by non-negative integers up to the population size N .
2. Using an indicator function $I(\cdot)$, the pmf of X , as given by

$$P(X = x|k) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} I(\max\{0, n - (N - k)\} \leq \text{integer } x \leq \min\{k, n\}),$$

follows from the definition of the hypergeometric distribution (i.e., for W given in the preamble) with parameters $M_1 = k$, $M_2 = N - k$, $m = n$. Based on a given outcome $X = x \in \{0, \dots, n\}$, the likelihood function for $k \in \Theta_k$ is

$$L(k) = P(X = x|k) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} I(n - (N - k) \leq x) I(x \leq k) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} I(x \leq k \leq N - (n - x)),$$

which is positive only if $x \leq k \leq N - (n - x)$.

3. For k with $x \leq k < N - (n - x)$, both $L(k) > 0$ and $L(k + 1) > 0$ hold by Question 2 and, using $\binom{b+1}{a} = \binom{b}{a} \frac{b+1}{b+1-a}$ for generic integers $0 \leq a \leq b$, we have

$$\frac{L(k)}{L(k+1)} = \frac{\frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}}{\frac{\binom{k+1}{x} \binom{N-k-1}{n-x}}{\binom{N}{n}}} = \frac{\binom{k}{x} \binom{N-k-1}{n-x} \frac{N-k}{N-k-(n-x)}}{\binom{k}{x} \frac{k+1}{k+1-x} \binom{N-k-1}{n-x}} = \frac{(N-k)(k+1-x)}{(N-k-n+x)(k+1)}$$

4. Using Question 3 for k with $x \leq k < N - (n - x)$, note

$$\frac{L(k)}{L(k+1)} = \frac{(N-k)(k+1-x)}{(N-k-n+x)(k+1)} > 1$$

holds if and only if

$$(N-k)(k+1-x) > (N-k-n+x)(k+1)$$

holds (as all terms above are positive integers), which is algebraically equivalent to

$$-x(N-k) > -(n-x)(k+1)$$

or $kn > (N+1)x - n$ or

$$k > k^* \equiv \frac{(N+1)x}{n} - 1;$$

the same relations hold upon replacing “ $>$ ” above with $<$ or $=$.

5. The likelihood $L(k)$ is only positive for integer k with $x \leq k \leq N - (n - x)$ (by Question 2) and, since $k^* < x$, it holds by Question 4 that $L(k) > L(k + 1)$ for each k with $k^* < x \leq k < N - (n - x)$; that is, $L(k) > L(k + 1) > 0$ holds for each integer $k = x, \dots, N - (n - x) - 1$ so that $L(k)$ is positive and strictly decreasing as a function of integer $k \in [x, N - (n - x)]$.

6. We need to maximize the likelihood $L(k)$ for integer k with $x \leq k \leq N - (n - x)$ (as $L(k) = 0$ for any other $k \in \Theta_k$).
- When $x = 0$, then $k^* \equiv \frac{(N+1)x}{n} - 1 = -1 < x = 0$ so that $\hat{k}_{\text{MLE}} = 0$ by Question 5 (since the likelihood is strictly decreasing for integer k with $0 = x \leq k \leq N - (n - x) = N - n$ here).
 - If $x \leq k^* < N - (n - x)$ holds where k^* is an integer, then $L(k) > L(k + 1)$ holds for any integer $k \in (k^*, N - (n - x))$ by Question 4 so that $L(k)$ is strictly decreasing over integers $k \in [k^* + 1, N - (n - x)]$. Likewise, $L(k) < L(k + 1)$ holds for any integer $k \in [x, k^*]$ by Question 4 so that $L(k)$ is strictly increasing over integers $k \in [x, k^*]$. Finally, as k^* is an integer with $x \leq k^* < N - (n - x)$, then $k^* + 1$ is an integer with $L(k^* + 1) > 0$ and, by Question 4, we have $L(k^*) = L(k^* + 1)$. In summary, $L(k^*) = L(k^* + 1) > L(k)$ holds for any $k \notin \{k^*, k^* + 1\}$ so that $\hat{k}_{\text{MLE}} = k^*$ or $k^* + 1$.

7. Under the $\text{ABC}(a, b, N)$ prior for k (with some $a, b \geq 0$), we want to show that the posterior distribution for k has a pmf proportional to

$$\begin{aligned} & \binom{a+k}{a+x} \binom{b+N-k}{b+(n-x)} I(k \in \{x, \dots, N - (n - x)\}) \\ &= \frac{(a+k)!(b+N-k)!}{(a+x)!(k-x)!(b-(n-x))!(N-k-(n-x))!} I(k \in \{x, \dots, N - (n - x)\}), \end{aligned}$$

or, equivalently (by removing terms above not involving k), proportional to

$$\frac{(a+k)!(b+N-k)!}{(k-x)!(N-k-(n-x))!} I(k \in \{x, \dots, N - (n - x)\}),$$

where $I(\cdot)$ denotes the indicator function. From “the likelihood $L(k) \times$ the $\text{ABC}(a, b, N)$ prior for k ,” we have that the pmf of the posterior distribution for k is proportional to

$$\begin{aligned} & \binom{k}{x} \binom{N-k}{n-x} I(k \in \{x, \dots, N - (n - x)\}) \times \binom{a+k}{a} \binom{b+N-k}{b} I(k \in \{0, \dots, N\}) \\ &= \frac{k!}{x!(k-x)!} \frac{(N-k)!}{(n-x)!(N-k-(n-x))!} \frac{(a+k)!}{a!k!} \frac{(b+N-k)!}{(N-k)!b!} I(k \in \{x, \dots, N - (n - x)\}), \end{aligned}$$

which (upon cancellation of terms $k!$ and $(N - k)!$ and removing unimportant terms not involving k) is proportional to

$$\frac{1}{(k-x)!} \frac{1}{(N-k-(n-x))!} (a+k)!(b+N-k)! I(k \in \{x, \dots, N - (n - x)\}),$$

as required.

8. If the conditional distribution $T|p$ of T given p is Binomial($N - n, p$) with pmf $P(T = t|p) = \binom{N-n}{t} p^t (1-p)^{N-n-t}$, $t = 0, \dots, N - n$, and the marginal distribution of p is Beta($x + 1, n - x + 1$) with density function $f(p) = \frac{[x+(n-x)]!}{x!(n-x)!} p^x (1-p)^{n-x}$, $0 < p < 1$, then the marginal pmf for T is given by

$$\begin{aligned} P(T = t) &= \int_0^1 P(T = t|p) f(p) dp = \int_0^1 \binom{N-n}{t} p^t (1-p)^{N-n-t} \frac{[x+(n-x)]!}{x!(n-x)!} p^x (1-p)^{n-x} dp \\ &= \binom{N-n}{t} \frac{[x+(n-x)]!}{x!(n-x)!} \int_0^1 p^{x+t} (1-p)^{N-x-t} dp \\ &= \binom{N-n}{t} \frac{[x+(n-x)]!}{x!(n-x)!} \frac{(x+t)!(N-x-t)!}{N!} \end{aligned}$$

for $t = 0, \dots, N - n$. Removing factorial terms not involving t above, $P(T = t)$ is proportional to

$$\frac{(x+t)!(N-x-t)!}{t!(N-n-t)!} I(t \in \{0, \dots, N - n\}).$$

Now if the posterior distribution for $k - x$ given $X = x \in \{0, \dots, n\}$ is ABC($x, (n-x), N - n$) (using $a = b = 0$), then the posterior probability $P(k - x = t|x)$ is proportional to

$$\begin{aligned} &\binom{x+t}{x} \binom{(n-x)+N-n-t}{(n-x)} I(t \in \{0, \dots, N - n\}) \\ &\propto \frac{(x+t)!}{t!} \frac{(N-x-t)!}{(N-n-t)!} I(t \in \{0, \dots, N - n\}) \\ &\propto P(T = t), \quad t = 0, \dots, N - n. \end{aligned}$$

The marginal distribution of T matches the posterior distribution of $k - x$, with a common support of $\{0, \dots, N - n\}$.

9. Under squared error loss, the Bayes estimator of k is the posterior mean of k . From Question 7, the posterior distribution of $k - x$ is ABC($x, (n-x), N - n$) which is the same as the marginal distribution of T from (T, p) in Question 8. Hence, the posterior mean of $k - x$ is the same as

$$ET = E(E[T|p]) = E((N-n)p) = (N-n)E(p) = (N-n)\frac{x+1}{n+2},$$

where $E[T|p] = (N-n)p$ follows from $T|p \sim \text{Binomial}(N - n, p)$ and $E(p) = \frac{x+1}{n+2}$ follows from the Beta($x + 1, n - x + 1$) mean of p given in Question 8. Consequently, the posterior mean of $k = (k - x) + x$ is

$$ET + x = (N-n)\frac{x+1}{n+2} + x = \frac{N-n+(N+2)x}{n+2} = \frac{(N+2)(x+1)}{n+2} - 1$$

for a given outcome $X = x \in \{0, \dots, n\}$.

Part II

- 10.** The distribution of X , for a given parameter $k \in \Theta_k \equiv \{0, \dots, N\}$, is Binomial($n, k/N$) with pmf

$$P(X = x|k) = \binom{n}{x} \left(\frac{k}{N}\right)^x \left(1 - \frac{k}{N}\right)^{n-x}, \quad x = 0, \dots, n.$$

For all $x \in \mathbb{R}$ and all $k \in \Theta_k$ and using an indicator function $I(\cdot)$, the probability

$$P(X = x|k) = \binom{n}{x} \left(\frac{k}{N}\right)^x \left(1 - \frac{k}{N}\right)^{n-x} I(x \in \{0, \dots, n\})$$

is a function of the data and unknown parameter k only through $x = \hat{k}_{\text{MOM}}n/N$ so that \hat{k}_{MOM} is sufficient for k by the factorization theorem.

- 11.** The MOM estimator of the parametric function $k(N - k)$ is given by $\hat{k}_{\text{MOM}}(N - \hat{k}_{\text{MOM}})$ for $\hat{k}_{\text{MOM}} = NX/n$. Using $E\hat{k}_{\text{MOM}} = (N/n)[EX] = (N/n)[nk/N] = k$ and

$$\begin{aligned} E[\hat{k}_{\text{MOM}}]^2 &= \frac{N^2}{n^2}[EX^2] = \frac{N^2}{n^2}(\text{Var}(X) + [EX]^2) \\ &= \frac{N^2}{n^2}(n(k)(N - k)/N^2 + [nk/N]^2) = \frac{k(N - k)}{n} + k^2, \end{aligned}$$

we have

$$\begin{aligned} E\hat{k}_{\text{MOM}}(N - \hat{k}_{\text{MOM}}) &= NE\hat{k}_{\text{MOM}} - E[\hat{k}_{\text{MOM}}]^2 = Nk - \frac{k(N - k)}{n} - k^2 \\ &= k(N - k) - \frac{k(N - k)}{n} = \frac{n - 1}{n}k(N - k). \end{aligned}$$

An unbiased estimator of $k(N - k)$ is then $n/(n - 1) \cdot \hat{k}_{\text{MOM}}(N - \hat{k}_{\text{MOM}})$.

- 12.** For $k_1 > k_0 = 0$ with $k_1 \in \{1, \dots, N\}$, we have

$$\frac{L(k_1)}{L(k_0)} = \left(\frac{k_1}{k_0}\right)^x \left(\frac{N - k_1}{N - k_0}\right)^{n-x} = \left(\frac{k_1}{0}\right)^x \left(\frac{N - k_1}{N}\right)^{n-x}, \quad x \in \{0, \dots, n\},$$

where we interpret $k_1/0 = +\infty$ for $k_1 > 0$ with $c \cdot (+\infty) = +\infty$ for any $c > 0$ and $c \cdot (+\infty) = 0$ when $c = 0$. If $k_1 < N$ then $N - k_1 > 0$ holds and $L(k_1)/L(0)$ is a non-decreasing function of x , being $0 < [(N - k_1)/N]^n \in \mathbb{R}$ when $x = 0$ and $(+\infty) \cdot [(N - k_1)/N]^{n-x} = +\infty$ for $x = 1, \dots, n$. If $k_0 = 0$ and $k_1 = N$, then $L(k_1)/L(k_0)$ is also a non-decreasing function of x , being 0 when $x = 0, \dots, n - 1$ and $(+\infty) \cdot [0/N]^0 = +\infty$ for $x = n$. Hence, for any $k_1 > k_0 = 0$ with $k_1 \in \{1, \dots, N\}$, $L(k_1)/L(k_0)$ is a nondecreasing function of x (or the statistic X).

- 13.** Since a monotone likelihood ratio property for $L(k_1/k_0)$ holds in X when $k_0 = 0$ and $k_1 > 0$ (i.e., $L(k_1/k_0)$ is nondecreasing in X), the UMP test of size $\alpha \in (0, 1)$ for $H_0: k = 0$ vs. $H_1: k > 0$ has the form

$$\varphi(X) = \begin{cases} 1 & \text{if } X > c \\ \gamma & \text{if } X = c \\ 0 & \text{if } X < c \end{cases}$$

with rejection occurring for large X values, where $\gamma \in [0, 1]$ and $c \in \mathbb{R}$ above are chosen so that

$$E_{k=0}\varphi(X) = P_{k=0}(X > c) + \gamma P_{k=0}(X = c) = \alpha \in (0, 1).$$

When $k = 0$, then $X = 0$ holds with probability 1 so that $P_{k=0}(X > c) = 1$ follows for values $c < 0$ and $P_{k=0}(X > c) = 0$ holds if $c \geq 0$. To get a size $\alpha \in (0, 1)$, one requires then $c \geq 0$ (to prevent the contradiction $\alpha = E_{k=0}\varphi(X) \geq P_{k=0}(X > c) = 1 > \alpha$) so that

$$E_{k=0}\varphi(X) = \gamma P_{k=0}(X = c) = \alpha \in (0, 1)$$

holds. But then one requires $c = 0$ to prevent $P_{k=0}(X = c) = 0$ when $c > 0$. Consequently, $c = 0$ with $\gamma = \alpha \in (0, 1)$ must hold in the form of the above test.

- 14.** The power of the size $\alpha \in (0, 1)$ test is then

$$E_k\varphi(X) = P_k(X > 0) + \alpha P_k(X = 0) = 1 - P_k(X = 0) + \alpha P_k(X = 0) = 1 - (1 - \alpha)P_k(X = 0)$$

where $P_k(X = 0) = (1 - k/N)^n$ so that the type II error probability for a given integer $k = 1, \dots, N$ is

$$1 - E_k\varphi(X) = (1 - \alpha)(1 - k/N)^n,$$

which is decreasing over $k = 1, \dots, N$ and hence bounded above by $(1 - \alpha)(1 - 1/N)^n$.

- 15.** Noting that $\hat{k}_{\text{MOM}} = NX/n$ and $X \stackrel{d}{=} \sum_{i=1}^n X_i$ for iid $X_1, \dots, X_n \sim \text{Bernoulli}(k/N)$, we have $\hat{k}_{\text{MOM}} \stackrel{d}{=} N\bar{X}_n$ for $\bar{X}_n = \sum_{i=1}^n X_i/n$. By the WLLN, $\bar{X}_n \xrightarrow{p} EX_1 = k/N$ as $n \rightarrow \infty$ so that $\sqrt{N - \hat{k}_{\text{MOM}}} \stackrel{d}{=} \sqrt{N - N\bar{X}_n} \xrightarrow{p} \sqrt{N - k}$ by the continuous mapping theorem. Additionally,

$$\sqrt{n}(\bar{X}_n - k/N) = \sqrt{n}(\bar{X}_n - EX_1) \xrightarrow{d} \text{Normal}\left(0, \text{Var}(X_1) = \frac{k}{N} \left[1 - \frac{k}{N}\right]\right)$$

as $n \rightarrow \infty$ by the CLT so that

$$\sqrt{n}(\hat{k}_{\text{MOM}} - k) \stackrel{d}{=} N\sqrt{n}(\bar{X}_n - k/N) \xrightarrow{d} \text{Normal}\left(0, N^2 \frac{k}{N} \left[1 - \frac{k}{N}\right]\right) \equiv \text{Normal}(0, k(N - k))$$

by Slutsky's theorem. Furthermore, as $\sqrt{N - \hat{k}_{\text{MOM}}} \xrightarrow{p} \sqrt{N - k}$ and $N - k > 0$ for $k < N$, we have

$$\frac{\sqrt{n}(\hat{k}_{\text{MOM}} - k)}{\sqrt{N - \hat{k}_{\text{MOM}}}} \xrightarrow{d} \frac{1}{\sqrt{N - k}} \cdot \text{Normal}(0, k(N - k)) \equiv \text{Normal}\left(0, \frac{k(N - k)}{N - k}\right) \equiv N(0, k)$$

as $n \rightarrow \infty$ by Slutsky's theorem.

- 16.** For a symmetric-type 95% confidence interval for k based on a distributional approximation of $\sqrt{n}(\hat{k}_{\text{MOM}} - k)$ with a normal with mean 0 and variance $k(N - \hat{k}_{\text{MOM}})$, we may use the set

$$\begin{aligned}
 & \left\{ k \in \Theta_k : \left| \sqrt{n}(\hat{k}_{\text{MOM}} - k) \right| \leq 1.96 \sqrt{k(N - \hat{k}_{\text{MOM}})} \right\} \\
 = & \left\{ k \in \Theta_k : n(\hat{k}_{\text{MOM}} - k)^2 \leq 1.96^2 k(N - \hat{k}_{\text{MOM}}) \right\} \\
 = & \left\{ k \in \Theta_k : k^2 - 2k\hat{k}_{\text{MOM}} + (\hat{k}_{\text{MOM}})^2 \leq n^{-1} 1.96^2 k(N - \hat{k}_{\text{MOM}}) \right\} \\
 = & \left\{ k \in \Theta_k : k^2 - 2kA_n + (\hat{k}_{\text{MOM}})^2 \leq 0 \right\} \\
 = & \left\{ k \in \Theta_k : A_n \pm \sqrt{A_n^2 - (\hat{k}_{\text{MOM}})^2} \right\}
 \end{aligned}$$

where $A_n = \hat{k}_{\text{MOM}} + 1.96^2(N - \hat{k}_{\text{MOM}})/(2n)$.

Part I

Let Ω be a non-empty set.

1. Define an algebra of subsets of Ω .
2. Define a σ -algebra of subsets of Ω .
3. Let \mathcal{C} be a collection of subsets of Ω . Define the σ -algebra generated by \mathcal{C} (denoted as $\sigma(\mathcal{C})$).
4. Let (Ω, \mathcal{F}) be a measurable space. Define a probability measure on (Ω, \mathcal{F}) .
5. Let B and C be two non-empty subsets of Ω . Define a sequence of sets $A_n \subset \Omega, n \geq 1$ such that

$$A_n = \begin{cases} B & \text{if } n \text{ is odd,} \\ C & \text{if } n \text{ is even.} \end{cases}$$

Find

$$\limsup_{n \rightarrow \infty} A_n \text{ and } \liminf_{n \rightarrow \infty} A_n.$$

6. Let \mathcal{D} be an algebra of subsets of Ω . Define

$$\bar{\mathcal{D}} = \{D \subset \Omega : \exists D_n \in \mathcal{D} \text{ for } n \geq 1 \text{ with } \lim_{n \rightarrow \infty} D_n = D\}.$$

- a) Show that $\mathcal{D} \subset \bar{\mathcal{D}}$.
- b) Show that $\bar{\mathcal{D}}$ is an algebra.

7. Let \mathcal{C} be the collection of all one point subsets (singletons) of Ω . Show that

$$\sigma(\mathcal{C}) = \{A \subset \Omega : A \text{ is countable}\} \cup \{A \subset \Omega : A^c \text{ is countable}\}.$$

8. Let (Ω, \mathcal{F}) be a measurable space. Let $\mu : \mathcal{F} \rightarrow [0, 1]$ be a set function satisfying the conditions,

- (i) μ is finitely additive on \mathcal{F} ,
- (ii) $\mu(\Omega) = 1$, and
- (iii) if $E_n \in \mathcal{F}, n \geq 1$ are disjoint with $\bigcup_{n=1}^{\infty} E_n = \Omega$, then $\sum_{n=1}^{\infty} \mu(E_n) = 1$.

Show that μ is a probability measure on (Ω, \mathcal{F}) .

9. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $G_n, n \geq 1$ be a monotone decreasing sequence of subsets of Ω . Is

$$\mu\left(\lim_{n \rightarrow \infty} G_n\right) = \lim_{n \rightarrow \infty} \mu(G_n)$$

always true? Prove your answer is correct.

Part II

Consider the Borel σ -algebra \mathcal{F} on $[0, 1]$, and the standard Borel space $([0, 1], \mathcal{F}, \lambda)$, where λ is the Lebesgue measure on \mathcal{F} . Define $X_n, n \geq 3$ on $([0, 1], \mathcal{F})$ where

$$X_n(\omega) = \frac{n}{\log n} I_{(0, \frac{1}{n}]}(\omega).$$

- 10.** Show that $X_n \rightarrow 0$ almost surely as $n \rightarrow \infty$.
- 11.** Show that $E(X_n) \rightarrow 0$ as $n \rightarrow \infty$.
- 12.** Does the condition of domination in the Dominated Convergence Theorem hold? That is, does there exist Y on $([0, 1], \mathcal{F}, \lambda)$ with $E|Y| < \infty$, and $|X_n| \leq Y$ for all $n \geq 3$?

Part III

Let $\{X_n, n \geq 1\}$ be random variables on a probability space (Ω, \mathcal{F}, P) with

$$P(X_n = n^2) = \frac{1}{n^2} \text{ and } P(X_n = -1) = 1 - \frac{1}{n^2}$$

for all $n \geq 1$. Let $S_n = \sum_{j=1}^n X_j$ for $n \geq 1$.

- 13.** Show that $\lim_{n \rightarrow \infty} S_n = -\infty$ almost surely.

Part IV

Let $\{X_n, n \geq 1\}$ be independent random variables on a probability space (Ω, \mathcal{F}, P) with

$$P(X_n = n^\alpha) = \frac{1}{2n^{2\alpha-1}} = P(X_n = -n^\alpha) \text{ and } P(X_n = 0) = 1 - \frac{1}{n^{2\alpha-1}}$$

for all $n \geq 1$, and for some $\alpha \geq 1/2$. Let $S_n = \sum_{j=1}^n X_j$ and $v_n^2 = \text{Var}(S_n)$, for $n \geq 1$.

- 14.** Show that for $\alpha \in [0.5, 1)$, as $n \rightarrow \infty$

$$\frac{S_n}{v_n} \xrightarrow{d} N(0, 1).$$

- 15.** Show that if $\alpha = 1$, for any given $\epsilon > 0$

$$\max_{1 \leq j \leq n} P(|X_j| > n\epsilon) \rightarrow 0 \text{ as } n \rightarrow \infty$$

- 16.** Show that for $\alpha > 1$, S_n converges almost surely as $n \rightarrow \infty$.

- 17.** For $\alpha > 1$, what happens to S_n/v_n as $n \rightarrow \infty$?

1. A non empty class \mathcal{C} of subsets of Ω is called an algebra if (i) $\Omega \in \mathcal{C}$, (ii) $A \in \mathcal{C} \Rightarrow A^c \in \mathcal{C}$, (iii) $A, B \in \mathcal{C} \Rightarrow A \cup B \in \mathcal{C}$.
2. A non empty class \mathcal{F} of subsets of Ω is called a σ -algebra if it is an algebra and if it satisfies (iv) $A_n \in \mathcal{F}, n \geq 1$ implies that $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.
3. The σ -algebra generated by \mathcal{C} , denoted as $\sigma(\mathcal{C})$ is a σ -algebra satisfying (i) $\mathcal{C} \subset \sigma(\mathcal{C})$, and (ii) if \mathcal{F} is another σ -algebra containing \mathcal{C} , then $\sigma(\mathcal{C}) \subset \mathcal{F}$.
4. A set function μ on \mathcal{F} is called a probability measure if (i) $\mu(A) \in [0, 1]$ for all $A \in \mathcal{F}$, (ii) $\mu(\Omega) = 1$, and (iii) for any disjoint collection of sets $A_n \in \mathcal{F}, n \geq 1$ with $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$,

$$\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n).$$

5.

$$\limsup_{n \rightarrow \infty} A_n = B \cup C \text{ and } \liminf_{n \rightarrow \infty} A_n = B \cap C.$$

6. a) Let $A \in \mathcal{D}$. Take $A_n = A$ for $n \geq 1$. Then $\lim_{n \rightarrow \infty} A_n = A$. Thus $\mathcal{D} \subset \bar{\mathcal{D}}$.
- b) (i) Since $\Omega \in \mathcal{D}$, and $\mathcal{D} \subset \bar{\mathcal{D}}$, $\Omega \in \bar{\mathcal{D}}$. (ii) Let $A \in \bar{\mathcal{D}}$. Thus, $\exists A_n \in \mathcal{D}, n \geq 1$, such that

$$A = \lim_{n \rightarrow \infty} A_n = \liminf_{n \rightarrow \infty} A_n = \cup_{k=1}^{\infty} \cap_{n=k}^{\infty} A_n = \cap_{k=1}^{\infty} \cup_{n=k}^{\infty} A_n = \limsup_{n \rightarrow \infty} A_n.$$

Since \mathcal{D} is an algebra, $A_n^c \in \mathcal{D}$. Also

$$A^c = \cap_{k=1}^{\infty} \cup_{n=k}^{\infty} A_n^c = \limsup_{n \rightarrow \infty} A_n^c = \liminf_{n \rightarrow \infty} A_n^c = \lim_{n \rightarrow \infty} A_n^c.$$

Thus $A^c \in \bar{\mathcal{D}}$. (iii) Let $A, B \in \bar{\mathcal{D}}$. Thus, $\exists A_n, B_n \in \mathcal{D}, n \geq 1$, such that

$$A = \lim_{n \rightarrow \infty} A_n, \quad B = \lim_{n \rightarrow \infty} B_n.$$

Since \mathcal{D} is an algebra, $A_n \cup B_n \in \mathcal{D}, n \geq 1$. We show that $A \cup B = \lim_{n \rightarrow \infty} A_n \cup B_n$. Thus $A \cup B \in \bar{\mathcal{D}}$.

Since $\limsup_{n \rightarrow \infty} A_n \cup B_n$ is the set of elements which are in $A_n \cup B_n$ for infinitely many n ,

$$\limsup_{n \rightarrow \infty} A_n \cup B_n = \limsup_{n \rightarrow \infty} A_n \cup \limsup_{n \rightarrow \infty} B_n = A \cup B.$$

Now

$$\liminf_{n \rightarrow \infty} A_n \cup B_n = \cup_{k=1}^{\infty} \cap_{n=k}^{\infty} (A_n \cup B_n) \supset \cup_{k=1}^{\infty} \cap_{n=k}^{\infty} A_n = \liminf_{n \rightarrow \infty} A_n = A.$$

Similarly

$$\liminf_{n \rightarrow \infty} A_n \cup B_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} (A_n \cup B_n) \supset \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} B_n = \liminf_{n \rightarrow \infty} B_n = B.$$

Thus

$$\liminf_{n \rightarrow \infty} A_n \cup B_n \supset A \cup B.$$

So

$$\limsup_{n \rightarrow \infty} A_n \cup B_n = \liminf_{n \rightarrow \infty} A_n \cup B_n = \lim_{n \rightarrow \infty} A_n \cup B_n = A \cup B.$$

Hence $\bar{\mathcal{D}}$ is an algebra.

- 7.** First we show that $\sigma(\mathcal{C})$ is a σ -algebra. First $\Omega \in \sigma(\mathcal{C})$ as $\Omega^c = \emptyset \in \sigma(\mathcal{C})$. By the definition of $\sigma(\mathcal{C})$, if $A \in \sigma(\mathcal{C})$, then $A^c \in \sigma(\mathcal{C})$. Let $A_n \in \sigma(\mathcal{C}), n \geq 1$. If A_n 's are countable for all $n \geq 1$, then $\bigcup_{n=1}^{\infty} A_n$ is countable and hence $\bigcup_{n=1}^{\infty} A_n \in \sigma(\mathcal{C})$. On the other hand, if for some n , A_n^c is countable, then $(\bigcup_{n=1}^{\infty} A_n)^c = \bigcap_{n=1}^{\infty} A_n^c$ is countable. Thus, $\bigcup_{n=1}^{\infty} A_n \in \sigma(\mathcal{C})$.

Since $\sigma(\mathcal{C})$ contains all countable subsets of Ω , it contains all singletons. Thus $\mathcal{C} \subset \sigma(\mathcal{C})$.

Finally, let \mathcal{F} be a σ -algebra containing \mathcal{C} . Let $A \in \sigma(\mathcal{C})$ and A be countable. Since A is a countable union of singletons, it must be in \mathcal{F} as \mathcal{F} is a σ -algebra containing \mathcal{C} . Now, let $A \in \sigma(\mathcal{C})$ and A^c be countable. Since A^c is a countable union of singletons, it is in \mathcal{F} . Since \mathcal{F} is a σ -algebra, $A = (A^c)^c \in \mathcal{F}$. Thus $\sigma(\mathcal{C}) \subset \mathcal{F}$.

- 8.** We need to show that μ is countably additive. Let $A_n \in \mathcal{F}, n \geq 1$ are disjoint. To show that $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$. By (i) and (ii),

$$1 = \mu(\Omega) = \mu((\bigcup_{n=1}^{\infty} A_n) \cup (\bigcup_{n=1}^{\infty} A_n)^c) = \mu(\bigcup_{n=1}^{\infty} A_n) + \mu((\bigcup_{n=1}^{\infty} A_n)^c). \quad (1)$$

By (ii) and (iii),

$$1 = \mu((\bigcup_{n=1}^{\infty} A_n) \cup (\bigcup_{n=1}^{\infty} A_n)^c) = \sum_{n=1}^{\infty} \mu(A_n) + \mu((\bigcup_{n=1}^{\infty} A_n)^c). \quad (2)$$

Combining (1) and (2), we have $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$.

- 9.** It is not always true. Let $\Omega = \mathbb{N}$, the set of natural numbers. Let μ be the counting measure, and $A_n = \{n, n+1, \dots\}$, for $n \geq 1$. Then $A_n \downarrow \emptyset$, $\mu(A_n) = \infty$, for all $n \geq 1$, but $\mu(\lim_{n \rightarrow \infty} A_n) = \mu(\emptyset) = 0$.

- 10.** Since for any $\omega > 0$, $X_n(\omega) = 0$ for all $n > 1/\omega$, $X_n \rightarrow 0$ almost surely (λ).

- 11.** $E(X_n) = 1/\log n \rightarrow 0$ as $n \rightarrow \infty$.

12. Since for $n \geq 3$, $n/\log n$ is increasing with n ,

$$\sup_{n \geq 3} X_n(\omega) = \sum_{n=3}^{\infty} \frac{n}{\log n} I_{(\frac{1}{n+1}, \frac{1}{n}]}(\omega).$$

Thus

$$E\left[\sup_{n \geq 3} X_n\right] = \sum_{n=3}^{\infty} \frac{n}{\log n} \frac{1}{n(n+1)} = \sum_{n=3}^{\infty} \frac{1}{(n+1)\log n} = \infty.$$

Hence the condition of domination in the DCT does not hold.

13. Since

$$\sum_{n=1}^{\infty} P(X_n = n^2) = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty,$$

by Borel Cantelli lemma, $P(X_n = n^2 \text{ infinitely often } (n)) = 0$. Thus, for almost all ω , $X_n(\omega) = -1$ eventually. So $\lim_{n \rightarrow \infty} S_n$ exists almost surely.

14. Also $\lim_{n \rightarrow \infty} S_n = -\infty$ almost surely.

15. Note that $EX_n = 0$, $EX_n^2 = n$. Thus $v_n^2 = \text{Var}(S_n) = \sum_{i=1}^n i = n(n+1)/2$. We now show that the Lindeberg Condition holds for $\alpha \in [0.5, 1)$. For a fixed $\epsilon > 0$,

$$\frac{1}{v_n^2} \sum_{j=1}^n EX_j^2 \mathbb{I}(|X_j| > \epsilon v_n) \leq \frac{1}{v_n^2} \sum_{j=1}^n EX_j^2 \mathbb{I}(|X_j| > \frac{\epsilon n}{\sqrt{2}}) = \frac{1}{v_n^2} \sum_{j=1}^n EX_j^2 \mathbb{I}(|X_j| > \frac{\epsilon n^\alpha n^{1-\alpha}}{\sqrt{2}})$$

Note that, for all $n > (\sqrt{2}/\epsilon)^{1/(1-\alpha)}$ and for all $j \leq n$, with probability 1,

$$\mathbb{I}(|X_j| > \frac{\epsilon n^\alpha n^{1-\alpha}}{\sqrt{2}}) = 0.$$

Thus

$$\frac{1}{v_n^2} \sum_{j=1}^n EX_j^2 \mathbb{I}(|X_j| > \epsilon v_n) \rightarrow 0$$

as $n \rightarrow \infty$. By Lindeberg CLT, $\frac{S_n}{v_n} \xrightarrow{d} N(0, 1)$.

16. Note for any given $\epsilon > 0$, for all $n > 1/\epsilon^2$, $n\epsilon > \sqrt{n}$. Thus for all $n > 1/\epsilon^2$ and $j < \sqrt{n}$, $P(|X_j| > n\epsilon) \leq P(|X_j| > \sqrt{n}) = 0$. So for all $n > 1/\epsilon^2$

$$\max_{1 \leq j \leq n} P(|X_j| > n\epsilon) = \max_{\lceil \sqrt{n} \rceil \leq j \leq n} P(|X_j| > n\epsilon) \leq \max_{\lceil \sqrt{n} \rceil \leq j \leq n} P(|X_j| \neq 0) = \max_{\lceil \sqrt{n} \rceil \leq j \leq n} \frac{1}{j} = \frac{1}{\lceil \sqrt{n} \rceil}.$$

Thus $\max_{1 \leq j \leq n} P(|X_j| > n\epsilon) \rightarrow 0$, as $n \rightarrow \infty$.

17. Take $c = 0.5$. For $\alpha > 1$,

$$\sum_{n=1}^{\infty} P(|X_n| > c) = \sum_{n=1}^{\infty} n^{1-2\alpha} < \infty.$$

Also for all n , $X_n^{(c)} = X_n I(|X_n| \leq c) = 0$ with probability 1. Thus $\sum_{n=1}^{\infty} E X_n^{(c)} = 0 = \sum_{n=1}^{\infty} \text{Var}(X_n^{(c)})$. So by Kolmogorov's 3-series theorem, S_n converges almost surely, as $n \rightarrow \infty$.

18. Since $v_n = \sqrt{n(n+1)/2} \uparrow \infty$, $S_n/v_n \rightarrow 0$ almost surely, as $n \rightarrow \infty$.

Part I

Let $\mathbf{X} \sim \mathcal{N}_n(\mathbf{0}, \mathbf{R})$, a multivariate normal of dimension n , where $\mathbf{X} = (X_1, \dots, X_n)'$ and \mathbf{R} is a correlation matrix with no negative elements and constant row sum. Let $\mathbf{R}^{\frac{1}{2}}$ be a square root matrix of \mathbf{R} which also has no negative entries. Further, assume that each row of $\mathbf{R}^{\frac{1}{2}}$ has the sum ϱ . Let $X_{(n)}$ be the maximum value of \mathbf{X} , that is, $X_{(n)} = \max\{X_1, X_2, \dots, X_n\}$.

1. Let $F_{(n)}(x) = P(X_{(n)} \leq x)$ be the cumulative distribution function (CDF) of $X_{(n)}$. Show that

$$F_{(n)}(x) \geq [\Phi(x/\varrho)]^n,$$

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

Part II

Let X be a realization from the normal distribution with zero mean and variance σ^2 , that is, $X \sim \mathcal{N}(0, \sigma^2)$. Let the CDF of X be denoted by $\Phi(x; \sigma)$ and the probability density function (PDF) of X be given by $\phi(x; \sigma)$. Let

$$W(x) = \frac{1 - \Phi(x; \sigma)}{\phi(x; \sigma)} \quad \text{and} \quad w(x) = \frac{d}{dx} W(x) \quad \text{for } x \in \mathbb{R}.$$

2. Show that $\lim_{x \rightarrow \infty} w(x) = 0$.
3. Using, for instance, the Mean Value Theorem, show that, for all $t \in \mathbb{R}$,

$$\frac{W(x + tW(x))}{W(x)} = 1 + tw\left(x \left[1 + \theta_{t,x} \frac{W(x)}{x}\right]\right), \quad \text{for some } 0 < \theta_{t,x} < 1.$$

4. From Problem 2 or directly, show that

$$\lim_{x \rightarrow \infty} xW(x) = \sigma^2.$$

5. The result from Problem 4 implies

$$\lim_{x \rightarrow \infty} \frac{W(x)}{x} = 0.$$

Using this result, show that, for each $t \in \mathbb{R}$,

$$\lim_{x \rightarrow \infty} \frac{W(x + tW(x))}{W(x)} = 1.$$

6. Let $x \equiv x_n = \Phi^{-1}(1 - 1/n; \sigma)$. Using the definition of $W(x)$ and the results in Problems 2-5, show that, for each $t \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} \Phi^n(x_n + tW(x_n); \sigma) = \exp\{-\exp(-t)\}.$$

Part III

The Department of the Prevention of Vice and the Protection of Virtue has only one candidate to lead it but who has a shortcoming of making (independent) random statements, some of which are false. Suppose that each statement made by the person is false with probability p . The candidate will not be further considered at the point that two consecutive false statements are heard. The Department begins monitoring the person's statements. Let X = the number of the statement at which the candidate ceases to be considered.

7. Show that $\mathbb{E}(X) = (1 + p)/p^2$. [Hint: Consider the four possible outcomes of the first two statements, namely “TT”, “TF”, “FT”, and “FF”. These each have corresponding probabilities and conditional means for X . Several of these conditional means are themselves simple functions of $\mathbb{E}(X)$, and an equation for $\mathbb{E}(X)$ in terms of conditional probabilities and means can be solved for in terms of p .]
8. Show that

$$\mathbb{E}(X^2) = \frac{2 - p^2 - p^3 + 4p}{p^4}$$

and

$$\text{Var}(X) = \frac{1 + 2p - 2p^2 - 2p^3}{p^4}.$$

9. Use the Markov inequality to produce inequalities for p that when satisfied guarantee that there is no more than a 1% chance of the candidate failing to be disqualified in the monitoring of only 100 statements. That is, based on the facts stated in Problems 7 and 8, for the positive random variables X and X^2 , give two inequalities for p so that when either is satisfied, we have $\mathbb{P}[X > 100] \leq 0.01$.

Part IV

Suppose X_1, X_2, \dots, X_n are iid $\text{Gamma}(\alpha_1, \beta)$, Y_1, Y_2, \dots, Y_n are iid $\text{Gamma}(\alpha_2, \beta)$, Z_1, Z_2, \dots, Z_n are iid $\text{Gamma}(\alpha_3, \beta)$, and all these random variables are independent.

10. What are the joint sufficient statistics for the parameters $(\alpha_1, \alpha_2, \alpha_3, \beta)$?
 11. If α_1, α_2 , and α_3 are known, derive the MLE for β .

Part I

Let $\mathbf{X} \sim \mathcal{N}_n(\mathbf{0}, \mathbf{R})$ where $\mathbf{X} = (X_1, \dots, X_n)'$ and \mathbf{R} is a correlation matrix with only non-negative elements such that $\mathbf{R}^{\frac{1}{2}}$, the square root matrix of \mathbf{R} , also has no negative entries. Further, assume that each row of \mathbf{R} has the same sum and so also does $\mathbf{R}^{\frac{1}{2}}$. Write $\mathbf{1} = (1, 1, \dots, 1)'$, we let ϱ be the sum of the elements in any row of $\mathbf{R}^{\frac{1}{2}}$. Further, let $X_{(n)}$ be the maximum value of \mathbf{X} , that is, $X_{(n)} = \max\{X_1, X_2, \dots, X_n\} \equiv \max \mathbf{X}$.

1. Show that the cumulative distribution function (CDF) $F_{(n)}(x)$ of $X_{(n)}$ is given by $F_{(n)}(x) = P(X_{(n)} \leq x) \geq [\Phi(x/\varrho)]^n$, where $\Phi(\cdot)$ is the CDF of the standard normal random variable.

- Let $\mathbf{Z} \sim \mathcal{N}_n(\mathbf{0}, \mathbf{I}_n)$ and $Z_{(n)} = \max\{Z_1, Z_2, \dots, Z_n\}$ have CDF $\Phi_{(n)}(z) \equiv [\Phi(z)]^n$. Then $\Phi_{(n)}(x/\varrho) = \mathbb{P}[Z_{(n)} \leq x/\varrho] = \mathbb{P}[\mathbf{Z} \leq x\mathbf{1}/\varrho]$ so that

$$\begin{aligned}\Phi_{(n)}(x/\varrho) &\leq \mathbb{P}[\mathbf{R}^{1/2}\mathbf{Z} \leq x\mathbf{R}^{1/2}\mathbf{1}/\varrho] \\ &= \mathbb{P}[\mathbf{X} \leq x\mathbf{1}], \text{ where } \mathbf{X} \sim N_n(\mathbf{0}, \mathbf{R}) \\ &= \mathbb{P}[X_{(n)} \leq x] = F_{(n)}(x).\end{aligned}\tag{1}$$

Part II

Let X be a realization from the normal distribution with zero mean and variance σ^2 , that is, $X \sim \mathcal{N}(0, \sigma^2)$. Let the cumulative distribution (CDF) of X be denoted by $\Phi(x; \sigma)$ and the probability density function (PDF) of X be given by $\phi(x; \sigma)$. Let

$$w(x) = \frac{1 - \Phi(x; \sigma)}{\phi(x; \sigma)}, \quad x \in \mathbb{R}$$

Answer the following questions:

2. Show that $\lim_{x \rightarrow \infty} \frac{d}{dx} w(x) = 0$.

- We have

$$\begin{aligned}\lim_{x \rightarrow \infty} \frac{d}{dx} w(x) &= \lim_{x \rightarrow \infty} \frac{-\phi(x; \sigma)^2 - \{\frac{d}{dx}\phi(x; \sigma)\}[1 - \Phi(x; \sigma)]}{\phi(x; \sigma)^2} \\ &= -1 - \lim_{x \rightarrow \infty} \frac{\frac{d}{dx}\phi(x; \sigma)}{\phi(x; \sigma)} \lim_{x \rightarrow \infty} \frac{1 - \Phi(x; \sigma)}{\phi(x; \sigma)} \\ &= -1 - \lim_{x \rightarrow \infty} \frac{\frac{d}{dx}\phi(x; \sigma)}{\phi(x; \sigma)} \lim_{x \rightarrow \infty} \frac{-\phi(x; \sigma)}{\frac{d}{dx}\phi(x; \sigma)}, \text{ by L'Hôpital's rule} \\ &= -1 + 1 = 0.\end{aligned}$$

3. Using, for instance, the Mean Value Theorem, show that

$$\frac{w(x + tw(x))}{w(x)} = 1 + tw' \left(x \left[1 + \theta_{t,x} \frac{w(x)}{x} \right] \right), \quad \text{for some } 0 < \theta_{t,x} < 1, \quad (2)$$

- As per the Mean Value Theorem, we have

$$w'(x + \theta_{t,x} w(x)) = \frac{w(x + tw(x)) - w(x)}{tw(x)}, \quad \text{for some } 0 < \theta_{t,x} < 1,$$

so that the result follows.

4. From Problem 2. or directly, show that

$$\lim_{x \rightarrow \infty} \frac{x}{w(x)} = \sigma^2. \quad (3)$$

and hence

$$\lim_{x \rightarrow \infty} \frac{w(x)}{x} = 0. \quad (4)$$

- From L'Hôpital's rule:

$$\lim_{x \rightarrow \infty} xw(x) = \lim_{x \rightarrow \infty} x \frac{-\phi(x; \sigma)}{\frac{d}{dx} \phi(x; \sigma)} = \lim_{x \rightarrow \infty} x \frac{1}{\frac{d}{dx} \log \phi(x; \sigma)} = \lim_{x \rightarrow \infty} x \frac{\sigma^2}{x} = \sigma^2,$$

and the result follows.

5. The result from Problem 4. implies

$$\lim_{x \rightarrow \infty} \frac{W(x)}{x} = 0.$$

Using this result, show that, for each $t \in \mathbb{R}$,

$$\lim_{x \rightarrow \infty} \frac{W(x + tW(x))}{W(x)} = 1.$$

- From (5.), we see that $1 + \theta_{t,x} \frac{w(x)}{x}$ is bounded away from 0, for all t and x . Then The right side of (2) converges uniformly to 1 in any bounded interval in \mathbb{R} .

6. Let $x \equiv x_n = \Phi^{-1}(1 - 1/n; \sigma)$. Using the definition of $w(x)$ and the results in Problems 2.-5., show that

$$\lim_{n \rightarrow \infty} \Phi^n(x_n + tw(x_n); \sigma) = \exp \{-\exp(-t)\}. \quad (5)$$

- We have

$$\lim_{n \rightarrow \infty} \frac{w(x_n + tw(x_n))}{w(x_n)} = 1.$$

Equivalently,

$$\lim_{n \rightarrow \infty} \frac{1 - \Phi(x_n + tw(x_n); \sigma)}{1 - \Phi(x_n; \sigma)} \frac{\phi(x_n; \sigma)}{\phi(x_n + tw(x_n); \sigma)} = 1$$

Or

$$\lim_{n \rightarrow \infty} n [1 - \Phi(x_n + tw(x_n); \sigma)] \frac{\phi(x_n; \sigma)}{\phi(x_n + tw(x_n); \sigma)} = 1$$

Now,

$$\lim_{n \rightarrow \infty} \frac{\phi(x_n; \sigma)}{\phi(x_n + tw(x_n); \sigma)} = \lim_{n \rightarrow \infty} \exp \left[\frac{tx_n w(x_n)}{\sigma^2} + \frac{t^2 w^2(x_n)}{2\sigma^2} \right] = \exp(t)$$

from (3) and (5.). This means that

$$\lim_{n \rightarrow \infty} n [1 - \Phi(x_n + tw(x_n); \sigma)] = \exp(-t)$$

Now,

$$\Phi^n(x_n + tw(x_n); \sigma) = \left\{ 1 - \frac{n [1 - \Phi(x_n + tw(x_n); \sigma)]}{n} \right\}^n$$

so that

$$\lim_{n \rightarrow \infty} \Phi^n(x_n + tw(x_n); \sigma) = \exp \{-\exp(-t)\}$$

Part III

The Department of the Prevention of Vice and the Protection of Virtue has only one candidate to lead it but who has a shortcoming of making (independent) random statements, some of which are false. Suppose that each statement made by the person is false with probability p . The candidate will not be further considered at the point that two consecutive false statements are heard. The Department begins monitoring the person's statements. Let

X = the number of the statement at which the candidate ceases to be considered.

7. Show that the expected number of statements (X) that the candidate makes before risking being disqualified is given by $\mathbb{E}(X) = (1 + p)/p^2$.

- Consider the three conditioning events: $A_1 = \{F\}$, $A_2 = \{SS\}$ and $A_3 = \{SF\}$, where these are the outcomes in the first and second statements. The A_i s are partitions of the statespace. Further,

$$\begin{aligned} \mu = \mathbb{E}(X) &= \mathbb{E}(X | A_1)\mathbb{P}(A_1) + \mathbb{E}(X | A_2)\mathbb{P}(A_2) + \mathbb{E}(X | A_3)\mathbb{P}(A_3) \\ &= (1 + \mu)(1 - p) + 2p^2 + (2 + \mu)p(1 - p) \end{aligned}$$

and the result follows.

8. Show that

$$\mu_2 = \mathbb{E}(X^2) = \frac{2 - p^2 - p^3 + 4p}{p^4}$$

and so the variance is given by

$$\text{Var}(X) = \frac{1 + 2p - 2p^2 - 2p^3}{p^4}.$$

- Proceeding in the same manner as 7., we have

$$\begin{aligned}\mu_2 &= \mathbb{E}(X^2) = \mathbb{E}[(X + 1)^2 | A_1]\mathbb{P}(A_1) + \mathbb{E}(X^2 | A_2)\mathbb{P}(A_2) + \mathbb{E}[(X + 2)^2 | A_3]\mathbb{P}(A_3) \\ &= \frac{2}{p^2} - 1 - p + \frac{4}{p} + \mu_2(1 - p^2)\end{aligned}$$

so that $\mu_2 = \frac{2-p^2-p^3+4p}{p^4}$ and the results follow.

9. Use the Markov inequality to produce inequalities for p that when satisfied guarantee that there is no more than a 1% chance of the candidate failing to be disqualified in the monitoring of only 100 statements. That is, based on the facts stated in Problems 7. and 8., for the positive random variables X and X^2 , give two inequalities for p so that when either is satisfied $\mathbb{P}[X > 100] \leq 0.01$.

- We have

$$\mathbb{P}[X > 100] = \mathbb{P}[X \geq 101] \leq \frac{\mathbb{E}(X^2)}{101^2} = \frac{2 - p^2 - p^3 + 4p}{101^2 p^4} \leq 0.01.$$

which yields $p \approx 0.43$.

Part IV

Suppose X_1, X_2, \dots, X_n is a sample from the $\text{Gamma}(\alpha_1, \beta_1)$ distribution, Y_1, Y_2, \dots, Y_n is a sample from the $\text{Gamma}(\alpha_2, \beta_2)$ distribution, and Z_1, Z_2, \dots, Z_n is a sample from the $\text{Gamma}(\alpha_3, \beta_3)$ distribution, and that all are random samples.

10. Given the three samples above, what are the joint sufficient statistics for the parameters $(\alpha_1, \alpha_2, \alpha_3, \beta)$.

- The joint likelihood of *all* the data is

$$\begin{aligned}&L(\alpha_1, \alpha_2, \alpha_3, \beta; X'_i s, Y'_i s, Z'_i s) \\ &= \prod_{i=1}^n \frac{1}{\beta^{\alpha_1} \Gamma(\alpha_1)} x_i^{\alpha_1-1} e^{-x_i/\beta} \times \prod_{i=1}^n \frac{1}{\beta^{\alpha_2} \Gamma(\alpha_2)} y_i^{\alpha_2-1} e^{-y_i/\beta} \times \prod_{i=1}^n \frac{1}{\beta^{\alpha_3} \Gamma(\alpha_3)} z_i^{\alpha_3-1} e^{-z_i/\beta} \\ &= \frac{(\prod_{i=1}^n x_i)^{\alpha_1-1} (\prod_{i=1}^n y_i)^{\alpha_2-1} (\prod_{i=1}^n z_i)^{\alpha_3-1}}{\beta^{n(\alpha_1+\alpha_2+\alpha_3)} [\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)]^n} e^{-\sum_{i=1}^n (x_i+y_i+z_i)/\beta}\end{aligned}$$

Let

$$\begin{aligned} & g(x'_i s, y'_i s, z'_i s, \alpha_1, \alpha_2, \alpha_3, \beta, U_1, U_2, U_3, U_4) \\ &= \frac{U_1^{\alpha_1-1} U_2^{\alpha_2-1} U_3^{\alpha_3-1}}{\beta^{n(\alpha_1+\alpha_2+\alpha_3)} [\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)]^n} e^{-U_4/\beta} \end{aligned}$$

and $h(x'_i s, y'_i s, z'_i s) = 1\{x_{(1)} > 0\}1\{y_{(1)} > 0\}1\{z_{(1)} > 0\}$, then

$$U_1 = \prod_{i=1}^n x_i, U_2 = \prod_{i=1}^n y_i, U_3 = \prod_{i=1}^n z_i, \text{ and } U_4 = \sum_{i=1}^n (x_i + y_i + z_i)$$

are jointly sufficient for $\alpha_1, \alpha_2, \alpha_3$, and β .

11. If α_1, α_2 , and α_3 are known, what is the MLE for β ?

- From part (a), we can obtain the log likelihood, now just a function of β , as

$$\begin{aligned} & l(\beta; x'_i s, y'_i s, z'_i s) \\ &= -\frac{1}{\beta} \sum_{i=1}^n (x_i + y_i + z_i) - n(\alpha_1 + \alpha_2 + \alpha_3) \ln \beta + C, \end{aligned}$$

where C is a constant with respect to the unknown β . (In fact, up to the constant, this log likelihood is identical to that of $W_1, W_2, \dots, W_n \sim \text{Gamma}(\alpha_1 + \alpha_2 + \alpha_3, \beta)$, with $x_i + y_i + z_i$ replaced with w_i . The derivative with respect to β is

$$\frac{d}{d\beta} l(\beta; x'_i s, y'_i s, z'_i s) = \frac{n\bar{x} + n\bar{y} + n\bar{z}}{\beta^2} - \frac{n(\alpha_1 + \alpha_2 + \alpha_3)}{\beta}.$$

Set this to 0 and solve for the MLE

$$\hat{\beta} = \frac{\bar{x} + \bar{y} + \bar{z}}{\alpha_1 + \alpha_2 + \alpha_3}.$$

The above demonstrates that $\hat{\beta}$ is a stationary point of the log likelihood function, but technically we need to know that it is a global maximum. First, we show that $\hat{\beta}$ is a local maximum. Noting that $\bar{w} = \bar{x} + \bar{y} + \bar{z}$, because all the sample sizes are equal, and letting $\alpha = \alpha_1 + \alpha_2 + \alpha_3$, we have

$$\frac{d^2}{d\beta^2} l(\beta; \cdot) \Big|_{\beta=\hat{\beta}} = -\frac{2n\bar{w}}{\hat{\beta}^3} + \frac{n\alpha}{\hat{\beta}^2} = -n \left(2\frac{\bar{w}\alpha^3}{\bar{w}^3} - \frac{\alpha^3}{\bar{w}^2} \right) = -\frac{n\alpha^3}{\bar{w}^2} < 0,$$

because $\alpha > 0$ and $\bar{w} > 0$. It is a global maximum if $\lim_{\beta \rightarrow 0+} L(\beta; \cdot) = 0$ at the boundary, but

$$L(\beta; \cdot) \propto \frac{e^{-n\bar{w}/\beta}}{\beta^{n\alpha}} = \frac{x^{n\alpha}}{e^{n\bar{w}x}}$$

for $x = \frac{1}{\beta} \rightarrow \infty$. Applying L'Hôpital's rule $\lceil n\alpha \rceil$ times yields

$$\lim_{\beta \rightarrow 0+} L(\beta; \cdot) = \lim_{x \rightarrow \infty} \frac{n\alpha(n\alpha-1)\cdots(n\alpha-\lfloor n\alpha \rfloor)}{(n\bar{w})^{\lceil n\alpha \rceil} x^{\lceil n\alpha \rceil - n\alpha} e^{n\bar{w}x}} = 0.$$