Suppose that Z is a binary random variable with $\Pr(Z=1) = p = 1 - \Pr(Z=0)$. Let X be a random variable whose distribution depends on Z as follows: if $Z=1$, then $X \sim \text{Poi}(\lambda_1)$ and if $Z=0$, then $X \sim \text{Poi}(\lambda_0)$, where $\text{Poi}(\lambda)$ refers to the Poisson distribution with pdf $g(x|\lambda) = \lambda^x e^{-\lambda}/x!$ for $\lambda > 0$ and $x = 0, 1, 2, \ldots$.

1. Derive the conditional mean and the conditional variance of X given that $Z = 0$.
2. Find the joint distribution of X and Z, that is to say, find an expression for $\Pr(X = x, Z = z)$.
3. Show that the pdf for the marginal distribution of X is

$$f(x) = \begin{cases} (1-p)(\lambda_0^x e^{-\lambda_0}/x!) + p(\lambda_1^x e^{-\lambda_1}/x!) & x = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

   and provide a possible interpretation for this distribution.
4. Find the mean and variance of X.
5. Suppose that $x_1, \ldots, x_n$ are iid observations from the pdf in part 3.
   (a) Write down the likelihood function for $x_1, \ldots, x_n$.
   (b) Let

$$w_i = \frac{(p)(\lambda_1^{x_i} e^{-\lambda_1}/x_i!)}{(1-p)(\lambda_0^{x_i} e^{-\lambda_0}/x_i!) + (p)(\lambda_1^{x_i} e^{-\lambda_1}/x_i!)}$$

   Show that $w_i = \Pr(Z = 1|X = x_i)$.
   (c) Show that solving the equations that define the maximum likelihood estimates is equivalent to finding p, λ_0, λ_1 satisfying:

$$\begin{aligned} p &= \frac{\sum_{i=1}^n w_i}{n} \\ \lambda_1 &= \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} \\ \lambda_0 &= \frac{\sum_{i=1}^n (1-w_i) x_i}{\sum_{i=1}^n (1-w_i)} \end{aligned}$$

   (Note these must be solved numerically because w_i depends on the parameter values.)

1. Given that $Z = 0$ we have X is $\text{Poi}(\lambda_0)$.
$E(X|Z = 0) = \sum_{x=0}^{\infty} x\lambda_0^x e^{-\lambda_0}/x! = \lambda_0 \sum_{x=1}^{\infty} \lambda_0^{x-1} e^{-\lambda_0}/(x-1)! = \lambda_0$.
Using the same type of calculation we find that $E(X(X-1)|Z = 0) = \lambda_0^2$ from which we calculate $\text{Var}(X|Z = 0) = \lambda_0$.

2. There are several ways to write the joint distribution. Here is one:
$\Pr(X = x, Z = z) = (p\text{Poi}(x|\lambda_1))^z \left((1-p)\text{Poi}(x|\lambda_0)\right)^{(1-z)}$
where $\text{Poi}(x|\lambda)$ is the Poisson pdf.

3. The marginal distribution is obtained by summing over the distribution of Z. One possible interpretation is that this could be the distribution of counts in a population consisting of two subpopulations – each described by a separate Poisson distribution – with p indicating the relative sizes of the subpopulations.

4. The mean and variance can be computed directly from the marginal distribution. It is easier to note that $E(X) = E(E(X|Z)) = p\lambda_1 + (1-p)\lambda_0$ and $\text{Var}(X) = E(\text{Var}(X|Z)) + \text{Var}(E(X|Z)) = p\lambda_1 + (1-p)\lambda_0 + p(1-p)(\lambda_1 - \lambda_0)^2$

5. (a) The likelihood function is $L = \prod_{i=1}^{n}(p\text{Poi}(x_i|\lambda_1) + (1-p)\text{Poi}(x_i|\lambda_0)$.
(b) The form of w_i follows directly from the definition of conditional probability.
(c) These equations are found by setting first derivatives of the log-likelihood to zero and doing some algebra.

Let $X_1, \ldots, X_n$ be i.i.d. random variables each having the exponential distribution with parameter θ and probability density function

$$p(x|\theta) = \frac{1}{\theta}e^{-x/\theta} \qquad x > 0, \theta > 0$$

1. Find the mean and variance of X_i.
2. Show that the pdf of $W = X_1 + X_2$ is $h(w|\theta) = \frac{w}{\theta^2}e^{-w/\theta}$ for $w > 0$ and $\theta > 0$.
3. Find the maximum likelihood estimator of θ, based on $X_1, \ldots, X_n$.
4. Show that the maximum likelihood estimator is unbiased and find its variance.
5. Let $X_{max} = \max(X_1, \ldots, X_n)$. Show that X_{max} does not have a limiting distribution.
6. With X_{max} as defined in the previous part, find the limiting distribution of $Y = X_{max} - \theta \log n$.
7. Suppose that we are interested in estimating $g(\theta) = \Pr(X > x^*) = e^{-x^*/\theta}$.
   (a) One possible estimator is $g(\hat{\theta})$, where $\hat{\theta}$ is the mle for θ from part 3. Find the asymptotic variance of $g(\hat{\theta})$.
   (b) It turns out that $g(\hat{\theta})$ is not unbiased. Let $T = \frac{1}{n}\sum_i I_{[X_i > x^*]}$, where $I_{[X_i > x^*]}$ is defined to be equal to 1 if $X_i > x^*$ and 0 otherwise. Show that T is an unbiased estimator of $g(\theta)$.
   (c) Find the variance of the unbiased estimator in part (b).
   (d) Find a better unbiased estimator than the unbiased estimator in part (b).

1. $E(X) = \int \frac{x}{\theta} e^{-x/\theta} dx = \theta \int y^{2-1} e^{-y} dy = \theta\Gamma(2) = \theta$
   Using similar approach $E(X^2) = 2\theta^2$ and the variance is θ^2.

2. Reasonably straightforward transformation.

3. $L = \theta^{-n} \exp(-\sum_i X_i/\theta)$. Setting the first derivative of the log likelihood to zero yields $-\frac{n}{\theta} + \frac{\sum_i X_i}{\theta^2} = 0$ so that $\hat{\theta} = \bar{X}$. It is easy to verify that this value is a maximizer of the likelihood.

4. $E(\hat{\theta}) = E(\bar{X}) = EX_i = \theta$ and $\text{Var}(\hat{\theta}) = \text{Var}(\bar{X}) = \text{Var}(X_i)/n = \theta^2/n$

5. $\Pr(X_{max} < x) = \prod_i \Pr(X_i < x) = (1 - e^{-x/\theta})^n$. As $n \to \infty$ this converges to zero for every x. Hence there is no limiting distribution.

6. Repeating the above yields $F_Y(x) = \Pr(Y < x) = (1 - \frac{1}{n}e^{-x/\theta})^n$. As $n \to \infty$, we find $F_Y(x) \to e^{\exp(-x/\theta)}$ which is a cdf defined on the entire real line.

7. Estimating $g(\theta) = \Pr(X > x^*) = e^{-x^*/\theta}$

   (a) Using the delta method, $\text{Var}[g(\bar{X})] = g'(\theta)^2 \text{Var}[\bar{X}] = \frac{(x^*)^2}{n\theta^2} e^{-2x^*/\theta}$.

   (b) $E(I_{[X_i > x^*]}) = \Pr(X_i > x^*) = g(\theta)$. Thus the sample mean of the indicators is unbiased.

   (c) Since this is the mean of iid Bernoulli trials, the variance is $g(\theta)(1-g(\theta))/n$.

   (d) The goal here is to mention (for some credit) and execute (for full credit) the Rao-Blackwell Theorem. The sufficient statistic is the sum of the X_i's. Then $E(T|\sum_i X_i = s) = E(I_{[X_i > x^*]}|\sum_i X_i = s) = \left(1 - \frac{x^*}{n\bar{X}}\right)^{(n-1)}$ (skipped some steps there!) is still unbiased and has lower variance.

Let $X_1, X_2, \ldots$ be iid random variables with $X_1 \sim$ Binomial $(2, \frac{\theta}{1+\theta}), \theta > 0$.

1. Suppose we want to test the hypotheses

$$H_0 : \theta = 1 \quad \text{vs.} \quad H_1 : \theta = 2 \tag{1}$$

based on $X_1, \ldots, X_n$, $n \geq 2$.

(a) Define a most powerful (MP) test of size α for the testing problem (1).

(b) Using the Neyman-Pearson lemma, find a size $\alpha = (n+1)2^{-2n}$ MP test for the hypotheses (1). Find *explicitly* the constants specifying your MP test.

[Hint: Under $\theta = 1$, $\sum_{i=1}^{n} X_i \sim$ Binomial $(2n, 1/2)$.]

2. Next, suppose that we want to find a large sample confidence interval for θ.

(a) Define the Fisher information number about θ in X_1.

(b) Show that the MLE of θ, based on $X_1, \ldots, X_n$, is given by

$$\hat{\theta}_n = \bar{X}_n/(2 - \bar{X}_n), \quad \text{where} \quad \bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i.$$

(Note: You do *not* have to check the second derivative condition.)

(c) Show that $\sqrt{n}(\hat{\theta}_n - \theta) \to^d N(0, \frac{\theta(1+\theta)^2}{2})$.

(d) Find a large sample $(1-\alpha)100\%$ two-sided confidence interval for θ based on your answer to part (c).

Ph.D. Prelim Solution/ STAT 542 - 543 (III) / Spring 2001

1. (a) -

(b) The joint pmf of $X_1, \cdots, X_n$ is

$$f(\underline{x}; \theta) = \prod_{i=1}^{n} \binom{2}{x_i} \left(\frac{\theta}{1+\theta}\right)^{x_i} \left(1 - \theta/_{1+\theta}\right)^{2-x_i} \cdot I_{\{0,1,2\}}(x_i)$$

$$= \prod_{i=1}^{n} \left[I_{\{0,1,2\}}(x_i) \binom{2}{x_i} \right] \cdot \frac{\theta^{\Sigma x_i}}{(1+\theta)^{2n}}$$

A size α most powerful test for $H_0: \theta = 1$ vs. $H_1: \theta = 2$ is given by

$$\phi(\underline{x}) = \begin{cases} 1 & \text{if } f(\underline{x}; 2)/f(\underline{x}; 1) > K \\ \gamma & \qquad = \\ 0 & \qquad < \end{cases}$$

where $K \in (0, \infty)$ and $\gamma \in [0,1]$ are such that

$$E_{\theta=1}\, \phi(\underline{X}) = \alpha \longrightarrow ①$$

Now,

$$\frac{f(\underline{x};2)}{f(\underline{x};1)} \gtrless k$$

$$\Leftrightarrow \quad \frac{2^{\Sigma x_i}}{3^{2n}} \Big/ \frac{1}{2^{2n}} \gtrless k$$

$$\Leftrightarrow \quad 2^{\Sigma x_i} \gtrless K_1$$

$$\Leftrightarrow \quad \Sigma x_i \gtrless K_2 \quad \left(= \frac{\log K_1}{\log 2}\right)$$

Hence, a size α MP test is given by

$$\phi(\underline{x}) = \begin{cases} 1 & \Sigma x_i > K_2 \\ \gamma & = \\ 0 & < \end{cases} \longrightarrow ②$$

where K_2 and γ satisfy ①, viz.

$$\alpha = E_{\theta=1}\, \phi(x) = P_{\theta=1}(\Sigma X_i > K_2) + \gamma \cdot P(\Sigma X_i = K_2)$$

Note that under $\theta = 1$,

$$Y \equiv \sum_{i=1}^{n} X_i \sim \text{Bin}(2n, \tfrac{1}{2}).$$

$$\Rightarrow \quad P(Y = 2n) = \binom{2n}{2n} \cdot 2^{-2n} = 2^{-2n}$$

$$P(Y = 2n-1) = \binom{2n}{2n-1} 2^{-2n} = 2n \cdot 2^{-2n}$$

Take $k_2 = 2n-1$. Then,

$$\begin{aligned}
(1+n)2^{-2n} &= P(Y > k_2) + \gamma \cdot P(Y = k_2) \\
&= P(Y = 2n) + \gamma \cdot P(Y = 2n-1) \\
&= 2^{-2n} + \gamma \cdot 2n\, 2^{-2n}
\end{aligned}$$

$$\Leftrightarrow \quad n\,2^{-2n} = \gamma \cdot 2n \cdot 2^{-2n} \quad \Leftrightarrow \quad \gamma = \tfrac{1}{2}.$$

Hence, a size $\alpha = (n+1)2^{-2n}$ test for $H_0 : \theta$

$H_1 : \theta = 2$ is given by (2) with $k_2 = 2n-1$, γ:

(a) -

(b).

$$L(\theta) = f(\underset{\sim}{x};\theta) = \prod_{i=1}^{n}\left[I_{\{0,1,2\}}(x_i)\binom{2}{x_i}\right]\cdot\frac{\theta^{\Sigma x_i}}{(1+\theta)^{2n}}$$

now

$$\frac{d}{d\theta}\log L(\theta) = 0$$

$$\Longleftrightarrow \quad \frac{d}{d\theta}\left[\Sigma x_i \cdot \log\theta - 2n\log(1+\theta)\right] = 0$$

$$\Leftrightarrow \quad \Sigma x_i \cdot \frac{1}{\theta} - \frac{2n}{1+\theta} = 0$$

$$\Leftrightarrow \quad (1+\theta)(\Sigma x_i) - 2n\theta = 0$$

$$\Leftrightarrow \quad \Sigma x_i = \theta(2n - \Sigma x_i)$$

$$\Leftrightarrow \quad \hat{\theta}_n = \frac{\bar{x}_n}{2-\bar{x}_n}.$$

(c). By the CLT,

$$\sqrt{n}\left(\bar{X}_n - E_\theta X_1\right) \xrightarrow{d} N\left(0,\ Var_\theta(X_1)\right)$$

i.e. $$\sqrt{n}\left(\bar{X}_n - \frac{2\theta}{1+\theta}\right) \xrightarrow{d} N\left(0,\ \frac{2\theta}{(1+\theta)^2}\right)$$

Now, $$h(x) = \begin{cases} \dfrac{x}{2-x}, & x \neq 2 \\ 0 & \text{o.w.} \end{cases}$$

is continuously differentiable at

$x = \dfrac{2\theta}{1+\theta}$ ($\neq 2$ for all $\theta > 0$), with

$$h'\left(\frac{2\theta}{1+\theta}\right) = \left.\frac{1.(2-x) - x(-1)}{(2-x)^2}\right|_{\frac{2\theta}{1+\theta}} = \left.\frac{2}{(2-x)^2}\right|_{\frac{2\theta}{1+\theta}}$$

$$= \frac{(1+\theta)^2}{2}$$

Hence, by the Delta-method,

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) = \sqrt{n}\left(h(\bar{X}_n) - h\left(\frac{2\theta}{1+\theta}\right)\right) \xrightarrow{d} N\left(0,\ \frac{(1+\theta)^4}{\quad}\cdot\frac{2\theta}{\quad}\right)$$

i.e
$$\sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow{d} N\left(0,\ \frac{\theta(1+\theta)^2}{2}\right).$$

2(d). Since $\hat{\theta}_n \xrightarrow[p]{} \theta$,

$$\frac{\sqrt{n}\,(\hat{\theta}_n - \theta)}{\sqrt{\dfrac{\hat{\theta}_n(1+\hat{\theta}_n)^2}{2}}} \xrightarrow{d} N(0,1).$$

Let z_α denote the α-quantile of $N(0,1)$, $\alpha \in (0,1)$. Then, a $100(1-\alpha)\%$ two-sided large sample CI for θ is given by

$$\left[\hat{\theta}_n - z_{(1-\alpha/2)} \cdot \sqrt{\frac{\hat{\theta}_n(1+\hat{\theta}_n)^2}{2n}},\ \hat{\theta}_n - z_{\alpha/2} \cdot \sqrt{\frac{\hat{\theta}_n(1+\hat{\theta}_n)^2}{2n}}\right]$$