Suppose that Y is a discrete random variable with probability mass function (pmf)

$$f(y) = \begin{cases} \alpha p^y(1-p) & y = 1, 2, 3, \ldots \\ 1 - \alpha p & y = 0 \end{cases}$$

where $0 < p < 1$ and α is a constant.

(i) Find the restrictions on α that are required for this to be a valid probability mass function.

(ii) Show that the expected value of Y is $\alpha p/(1-p)$.
(For future reference the variance of Y is $(\alpha p + \alpha p^2 - \alpha^2 p^2)/(1-p)^2$.)

(iii) Find the conditional distribution of Y given that $Y \geq 1$.

(iv) Assume that $Y_1, \ldots, Y_n$ are n independent, identically distributed random variables having the pmf $f(y)$. Let $\bar{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i$. Identify the asymptotic distribution of $\bar{Y}_n$.

(v) Write down the likelihood function and identify a two-dimensional sufficient statistic.

(vi) The maximum likelihood estimates are given by:

$$\begin{aligned} \hat{p} &= 1 - \frac{n - N_o}{\sum_{i=1}^n Y_i}, \\ \hat{\alpha} &= \left(1 - \frac{N_o}{n}\right)/\hat{p}, \end{aligned}$$

where $N_o = \sum_{i=1}^n I_{\{0\}}(Y_i)$ and $I_A(X)$ is the indicator function that is one if $X \in A$ and zero otherwise.

(1) Find the maximum likelihood estimator $\hat{\theta}$ for the parameter $\theta = \alpha p$.

(2) Is $\hat{\theta}$ an unbiased estimator for the parameter θ?

(3) Find the variance of $\hat{\theta}$.

(vii) Describe and carry out one approach for obtaining the asymptotic mean and variance of $\hat{p}$ defined in (vi). Your result should be a function of n, α, p only but you need not simplify products and quotients.

Suppose that Y is a discrete random variable with probability mass function (pmf)

$$f(y) = \begin{cases} \alpha p^y(1-p) & y = 1, 2, 3, \dots \\ 1 - \alpha p & y = 0 \end{cases}$$

where $0 < p < 1$ and α is a constant.

(i) Note that $\sum_{y=0}^{\infty} f(y) = 1$ for any α as long as $0 < p < 1$. Also note that $f(y) \geq 0$ if $0 \leq \alpha \leq p^{-1}$.

(ii) $E(Y) = \alpha p(1-p)\sum_{y=1}^{\infty} y p^{y-1} = \alpha p(1-p)\frac{d}{dp}\left(\sum_{y=1}^{\infty} p^y\right) = \alpha p(1-p)\frac{d}{dp}\left(\frac{p}{1-p}\right) = \frac{\alpha p}{1-p}$.

(iii) $\Pr(Y = y|Y \geq 1) = \Pr(Y = y)/\Pr(Y \geq 1) = p^{y-1}(1-p)$ for $y = 1, 2, \dots$. (This is the pmf for a geometric distribution.)

(iv) Applying the CLT we have $\sqrt{n}(\bar{Y}_n - \frac{\alpha p}{1-p})$ converging in distribution to a $N(0, (\alpha p + \alpha p^2 - \alpha^2 p^2)/(1-p)^2)$ distribution. (Any such result is acceptable.)

(v) Need to introduce indicators for the values of Y. Let N_j represent the number of observations in the random sample for which $Y = j$. Then we find

$$\begin{aligned} L &= (1-\alpha p)^{N_o} \prod_{y=1}^{\infty} (\alpha p^y(1-p))^{N_y} \\ &= (1-\alpha p)^{N_o}(\alpha(1-p))^{\sum_{y=1}^{\infty} N_y} p^{\sum_{y=1}^{\infty} y N_y} \\ &= (1-\alpha p)^{N_o}(\alpha(1-p))^{n-N_o} p^{\sum_{i=1}^{n} Y_i}, \end{aligned}$$

from which it follows that N_o and $\sum_{i=1}^{n} Y_i$ are sufficient statistics.

(vi) (1) By the invariance principle $\hat{\theta} = \hat{\alpha}\hat{p} = 1 - (N_o/n)$.

(2) $E(\hat{\theta}) = 1 - E(N_o)/n = 1 - \Pr(Y = 0) = \alpha p = \theta$ so the estimator is unbiased.

(3) Note that the estimator is the complement of the proportion of the Y_i's that are equal to zero. Since the Y_i's are iid, the indicators are iid Bernoulli trials with probability of success $(1 - \alpha p)$. Thus $\text{Var}\ \hat{\theta} = (1-\alpha p)\alpha p/n$.

(vii) Hard – There are two approaches that work: rely on maximum likelihood theory to identify the asymptotic mean (p) and the asymptotic variance (inverse of the Fisher Information matrix); or rely on the asymptotic normality of N_o/n and $\bar{Y}_n$ and use the delta method. The second approach is now described in some detail.

Note that the mle $\hat{p}$ is a function of the sample mean $\bar{Y}_n$ and the sample proportion N_o/n. Let's call the sample proportion of zero values $\bar{I}_n$ as it is the mean of indicators. Then we can write

$$\hat{p} = g(\bar{I}_n, \bar{Y}_n) = 1 - \frac{1 - \bar{I}_n}{\bar{Y}_n}.$$

We can expand this bivariate function in a Taylor Series around the expected values, $E(\bar{Y}_n) = \alpha p/(1-p)$ and $E(\bar{I}_n) = 1 - \alpha p$, as follows

$$\begin{aligned}
\hat{p} &\approx p + (\bar{I}_n - (1-\alpha p)) \left.\frac{\partial g(\bar{I}_n, \bar{Y}_n)}{\partial \bar{I}_n}\right|_{(E\bar{I}_n, E\bar{Y}_n)} + (\bar{Y}_n - \frac{\alpha p}{1-p}) \left.\frac{\partial g(\bar{I}_n, \bar{Y}_n)}{\partial \bar{Y}_n}\right|_{(E\bar{I}_n, E\bar{Y}_n)} \\
&= p + (\bar{I}_n - (1-\alpha p)) \left(\frac{1-p}{\alpha p}\right) + (\bar{Y}_n - \frac{\alpha p}{1-p}) \left(\frac{(1-p)^2}{\alpha p}\right)
\end{aligned}$$

ignoring higher order terms (the remainder is $O(1/n)$ assuming $\alpha > 0$). Then the expected value of the right-hand-side is p. The variance of the right-hand-side is

$$\left(\frac{(1-p)^2}{(\alpha p)^2}\right) \text{Var } \bar{I}_n + \left(\frac{(1-p)^4}{(\alpha p)^2}\right) \text{Var } \bar{Y}_n + +2\left(\frac{(1-p)^3}{(\alpha p)^2}\right) \text{Cov}(\bar{I}_n, \bar{Y}_n).$$

The variance of $\bar{I}_n$ is available from (vi)(3) and the variance of $\bar{Y}_n$ is easily obtained from the information provided in (ii). Since the bivariate pairs I_i, Y_i are independent the covariance is just $\frac{1}{n}(E(I_iY_i) - E(I_i)E(Y_i))$; but the expectation of the product is zero since $I_i = 1$ only if $Y_i = 0$ and so the covariance is $-(1-\alpha p)\alpha p/(n(1-p))$. Combining these terms requires considerable algebra (the questions tells everyone not to bother) but yields the relatively simple answer $(1-p)^2/(\alpha n)$.

Stat 542-543 II

Throughout this question, for $\theta_1 > 0$ and $\theta_2 \in \mathcal{R}$, let $f(x|\theta_1,\theta_2)$ be the two parameter exponential probability density

$$f(x|\theta_1,\theta_2) = \begin{cases} \theta_1 \exp(-\theta_1(x-\theta_2)) & \text{for } x \geq \theta_2 \\ 0 & \text{otherwise} \end{cases} .$$

(a) Show that if $U \sim \text{Uniform}(0,1)$, then $Y = -\ln U$ has the "standard" exponential distribution (i.e. the exponential distribution with $\theta_1 = 1$ and $\theta_2 = 0$).

(b) Based on the fact in (a) and the availability of a uniform random number generator, how do you suggest simulating a random sample of size n from the Exponential(θ_1,θ_2) distribution? (Here we suppose that θ_1 and θ_2 are known constants.)

Let $X_1, X_2, \ldots, X_n$ be iid with marginal density $f(x|\theta_1,\theta_2)$.

(c) Find the (joint) maximum likelihood estimators of θ_1 and θ_2 based on $X_1, X_2, \ldots, X_n$.

(d) Argue carefully that your MLE of θ_1 from (c) is consistent.

(e) Give the form of the likelihood ratio tests of

i) $H_0\!:\theta_1 = \theta_1^0$ vs $H_a\!:$not H_0

and ii) $H_0\!:\theta_2 = \theta_2^0$ vs $H_a\!:$not H_0

(Write complete formulas for the test statistics and indicate what kinds of values of these will cause rejection of the null hypotheses. But you need NOT simplify the forms of your statistics nor speculate how to choose cut-offs to get α level tests.)

For each $i = 1, 2, \ldots, n$ let

$$X_i^* = X_i \text{ rounded to the nearest integer}$$

and now suppose that one gets to observe not the X_i's but rather the X_i^*'s. In fact, suppose that the frequency distribution of what is observed in a sample of $n = 20$ is

x^*	5	6	7	8	9
frequency	7	8	2	2	1

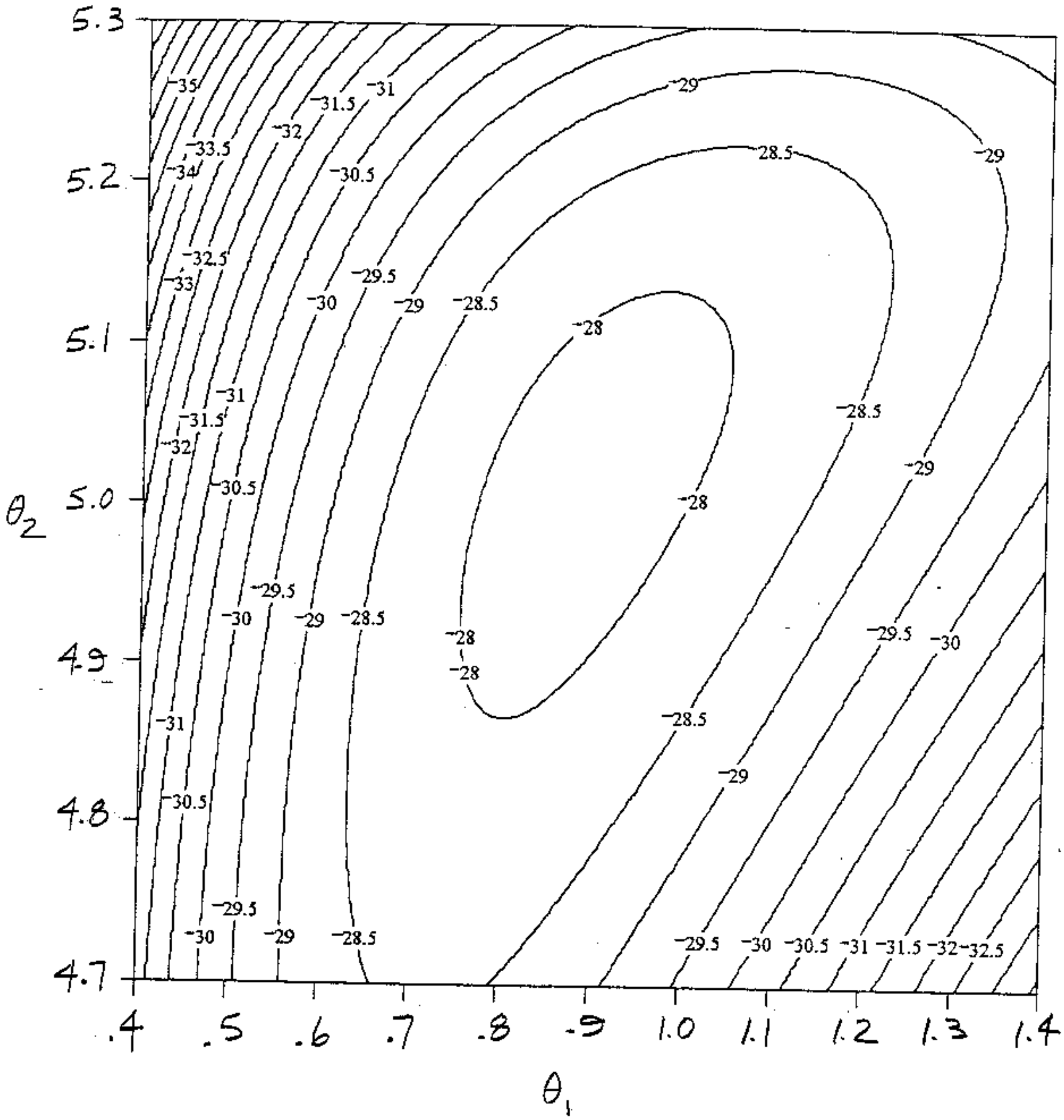
(f) Write out an appropriate likelihood function $L^*(\theta_1, \theta_2)$ based on these data. (Hint: For $x^* = 5, 6, 7, 8$ and 9, what are $P_{\theta_1,\theta_2}[X^* = x^*]$? Be careful. You will need one prescription for the likelihood in cases where $\theta_2 < 4.5$ and another for cases where $\theta_2 \geq 4.5$.)

As a matter of fact, the likelihood function referred to in (f) is maximized at $\theta_1 = .894$ and $\theta_2 = 5.018$. With $\mathcal{L}(\theta_1, \theta_2) = \ln L^*(\theta_1, \theta_2)$, $\mathcal{L}(.894, 5.018) = -27.833$ and second partial derivatives of $\mathcal{L}$ evaluated at $(.894, 5.018)$ are

$$\frac{\partial^2 \mathcal{L}}{\partial \theta_1^2} = -23.9, \quad \frac{\partial^2 \mathcal{L}}{\partial \theta_2^2} = -29.7 \text{ and } \frac{\partial^2 \mathcal{L}}{\partial \theta_1 \partial \theta_2} = 16.0 \ .$$

(g) Give maximum likelihood estimates of θ_1 and θ_2 and appropriate standard errors for the estimates. Do the data suggest that maximum likelihood estimators based on X_i^*'s will be positively correlated or negatively correlated? Explain.

(h) Attached to this question is a contour plot of $\mathcal{L}(\theta_1, \theta_2)$. Use it and test the hypothesis $\mathrm{H_0}: \theta_1 = 1.30$ and $\theta_2 = 4.90$ vs $\mathrm{H_a}$:not $\mathrm{H_0}$, with $\alpha \approx .05$. (There is a table of χ^2 percentage points also attached to this question.)


5.3
5.2
5.1
5.0
4.9
4.8
4.7
θ_2
.4 .5 .6 .7 .8 .9 1.0 1.1 1.2 1.3 1.4
θ_1
-35
-34
-33.5
-33
-32.5
-32
-31.5
-31
-30.5
-30
-29.5
-29
-28.5
-28
-28.5
-29
-29.5
-30
-30.5
-31
-31.5
-32
-32.5

Stat 542/543 II Page 4 of 4

Table A.7 Chi-Square Distribution Quantiles

ν	$Q(.005)$	$Q(.01)$	$Q(.025)$	$Q(.05)$	$Q(.1)$	$Q(.9)$	$Q(.95)$	$Q(.975)$	$Q(.99)$	$Q(.995)$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.143	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.290	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.653	40.647	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.994
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
31	14.458	15.655	17.539	19.281	21.434	41.422	44.985	48.232	52.192	55.003
32	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328
33	15.815	17.074	19.047	20.867	23.110	43.745	47.400	50.725	54.775	57.648
34	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964
35	17.192	18.509	20.569	22.465	24.797	46.059	49.802	53.204	57.342	60.275
36	17.887	19.233	21.336	23.269	25.643	47.212	50.998	54.437	58.619	61.581
37	18.586	19.960	22.106	24.075	26.492	48.364	52.192	55.668	59.893	62.885
38	19.289	20.691	22.878	24.884	27.343	49.513	53.384	56.896	61.163	64.183
39	19.996	21.426	23.654	25.695	28.196	50.660	54.572	58.120	62.429	65.477
40	20.707	22.164	24.433	26.509	29.051	51.805	55.759	59.342	63.691	66.767

This table was generated using the "INVCDF" command in Minitab.

For $\nu > 40$ the approximation $Q(p) \approx \nu\left(1 - \frac{2}{9\nu} + Q_z(p)\sqrt{\frac{2}{9\nu}}\right)^3$ can be used.

a) For $t>0$

$$P[Y>t] = P[-\ln U > t] = P[\ln U < -t] = P[U < e^{-t}]$$
$$= e^{-t}$$

b) Generate $U_1, U_2, \ldots, U_n$ iid Uniform$(0,1)$. Then take

$$X_i = \theta_1(-\ln U_i) + \theta_2 .$$

$X_1, X_2, \ldots, X_n$ are iid Exp(θ_1, θ_2)

c) The likelihood here is (for $\theta_1 > 0$ and $\theta_2 \leq \min X_i$)

$$L(\theta_1,\theta_2) = \prod_{i=1}^{n} f(X_i|\theta_1,\theta_2)$$
$$= \theta_1^n \exp -\theta_1 \sum_{i=1}^{n}(X_i - \theta_2)$$

So $\mathcal{L}(\theta_1,\theta_2) = \ln L(\theta_1,\theta_2) = n\ln\theta_1 - \theta_1 \sum_{i=1}^{n}(X_i-\theta_2)$

$\mathcal{L}(\theta_1,\theta_2)$ clearly increases in θ_2 for every θ_1, $\therefore \hat{\theta}_2 = \min X_i$

Then
$$\mathcal{L}(\theta_1,\hat{\theta}_2) = n\ln\theta_1 - \theta_1\sum_{i=1}^{n}(X_i - \text{mix } X_i)$$

is maximized at $\hat{\theta}_1 = \dfrac{n}{\sum(X_i - \min X_i)}$

So joint MLE's are $\hat{\theta}_1$ and $\hat{\theta}_2$.

d) Note that $\hat{\theta}_2 = \min X_i$ converges in θ_1,θ_2 probability to θ_2. Why? For $\epsilon > 0$

$$P_{\theta_1,\theta_2}[\hat{\theta}_2 > \theta_2 + \epsilon] = P_{\theta_1,\theta_2}[\min X_i > \theta_2+\epsilon]$$
$$= \left(P_{\theta_1,\theta_2}[X_1 > \theta_2 + \epsilon]\right)^n$$
$$\longrightarrow 0$$

Then $\hat{\theta}_1 = \dfrac{1}{\frac{1}{n}\sum(X_i-\theta_2) + \underbrace{(\theta_2 - \hat{\theta}_2)}_{\to 0}}$

and the LLN says that $\frac{1}{n}\sum(X_i-\theta_2) \xrightarrow[P_{\theta_1,\theta_2}]{} \frac{1}{\theta_1}$.
Since $g(a,b) = \frac{1}{a+b}$ is cont$^{\underline{s}}$ at $\left(\frac{1}{\theta_1}, 0\right)$ we're done.

e) i) Let
$$\lambda_1 = \frac{\sup L(\theta_1,\theta_2)}{\sup L(\theta_1^0,\theta_2)} = \frac{L(\hat{\theta}_1,\hat{\theta}_2)}{L(\theta_1^0,\hat{\theta}_2)}$$

and reject for large values of λ_1

ii) Let $\hat{\hat{\theta}}_1(\theta_2) = \dfrac{n}{\sum(X_i-\theta_2)}$. Take
$$\lambda_2 = \frac{\sup L(\theta_1,\theta_2)}{\sup L(\theta_1,\theta_2^0)} = \frac{L(\hat{\theta}_1,\hat{\theta}_2)}{L(\hat{\hat{\theta}}_1(\theta_2^0),\theta_2^0)}$$

and reject for large values of λ_2

f) For $\theta_2 \le 5.5$ and $x^* = 6, 7, 8$ or 9
$$P_{\theta_1,\theta_2}[X^* = x^*] = e^{-\theta_1(x^*-.5-\theta_2)} - e^{-\theta_1(x^*+.5-\theta_2)}$$

For $\theta_2 < 4.5$
$$P_{\theta_1,\theta_2}[X^* = 5] = e^{-\theta_1(4.5-\theta_2)} - e^{-\theta_1(5.5-\theta_2)}$$
and for $\theta_2 \in [4.5, 5.5]$
$$P_{\theta_1,\theta_2}[X^* = 5] = 1 - e^{-\theta_1(5.5-\theta_2)}$$

Finally, with these conventions, for $\theta_1 > 0$ and $\theta_2 < 5.5$
$$L(\theta_1,\theta_2) = \left(P_{\theta_1,\theta_2}[X^*=5]\right)^7 \left(P_{\theta_1,\theta_2}[X^*=6]\right)^8 \left(P_{\theta_1,\theta_2}[X^*=7]\right)^2$$
$$\times \left(P_{\theta_1,\theta_2}[X^*=8]\right)^2 P_{\theta_1,\theta_2}[X^*=9]$$

g) The observed Fisher information matrix is

$$\hat{I} = -\begin{bmatrix} -23.9 & 16.0 \\ 16.0 & -29.7 \end{bmatrix} \quad , \text{so}$$

$$\hat{I}^{-1} = \begin{bmatrix} .065 & .035 \\ .035 & .053 \end{bmatrix}$$ which is an estimated

variance-covariance matrix for the vector of MLE's.

$\hat{\theta}_1 = .894$ with std error $\sqrt{.065} = .25$
$\hat{\theta}_2 = 5.018$ with std error $\sqrt{.053} = .23$
and the estimated correlation is positive.

h) Large sample Theory says that under $\theta_1 = 1.30$ and $\theta_2 = 4.90$

$$2\left(\mathcal{L}(\hat{\theta}_1, \hat{\theta}_2) - \mathcal{L}(1.30, 4.90) \right) \xrightarrow{D} \chi^2_2$$

What has been observed is

$$2\left(\mathcal{L}(\hat{\theta}_1, \hat{\theta}_2) - \mathcal{L}(1.30, 4.90) \right) \approx 2\left(-27.833 - (-30.269)\right)$$

$$= 4.872$$

Now the upper .05 pt of the χ^2_2 dsn is 5.991.
So with $\alpha \approx .05$ accept H_0: $\theta_1 = 1.30$ and $\theta_2 = 4.90$

542–543 Question III Ph.D. Prelim Exam Spring 1999

Let X be an exponential random variable with mean λ. Let $M > 0$ and $X^{(M)}$ be the truncated random variable that is X if $X \le M$ and M if $X > M$.

1) Find the mean and variance of $X^{(M)}$.

2) Show that $X^{(M)}$ converges in probability and in distribution to X as $M \to \infty$.

3) Find the conditional distribution of $X - M$ given that $X^{(M)} = M$.

4) Find the best mean square predictor of X based on $X^{(M)}$. That is, a function $h(X^{(M)})$ such that $E(X-h(X^{(M)}))^2$ is minimized.

5) Let $X_1, X_2, \ldots, X_n$ be i.i.d. $\exp(\lambda)$ r.v. Let

$$Z_n^{(M)} = \frac{1}{n} \sum_{j=1}^{n} X_j^{(M)}$$

be the mean of the truncated r.v. where $X_j^{(M)} = X_j$ if $X_j \le M$ and M if $X_j > M$.

i) Show that $Z_n^{(M)}$ is a biased estimator of λ. Compute the bias and show that it goes to zero as $M \to \infty$.

ii) Show that $Z_n^{(M)}$ converges to λ in probability as n and M go to ∞.

iii) By comparing $Z_n^{(M)}$ with $\overline{X}_n = \frac{1}{n} \sum_{j=1}^{n} X_j$ guess the asymptotic distribution of $Z_n^{(M)}$ for large n and M stating the asymptotic mean and variance.

1) Since $X^{(M)} = X$ if $X \le M$ and M if $X > M$,
and X is an exponential r-v. with mean λ

$E X^{(M)} = \frac{1}{\lambda}\int_0^{\lambda} x e^{-x/\lambda} dx + M e^{-M/\lambda} = \lambda(1 - e^{-M/\lambda})$ and

$E(X^{(M)})^2 = \frac{1}{\lambda}\int_0^{\lambda} x^2 e^{-x/\lambda} dx + M^2 e^{-M/\lambda} = 2\lambda^2(1 - e^{-M/\lambda}) - 2\lambda e^{-M/\lambda}$

2) For $\epsilon > 0$, $|X^{(M)} - X| > \epsilon \implies X > M$ and so

$P(|X^{(M)} - X| > \epsilon) \le P(X > M) \le e^{-M/\lambda} \to 0$ as $M \to \infty$

3) For $t > 0$

$P(X - M > t \mid X^{(M)} = M) = P(X > M + t \mid X > M)$

$= P(X > M+t)/P(X > M) = e^{-\frac{(M+t)}{\lambda}} / e^{-M/\lambda} = e^{-t/\lambda}$

ie $X - M \mid X^{(M)} = M$ is ~ exp with mean λ.

4) It is known that the best mean square predictor given some information is the conditional expectation

Now if $X^{(M)} = M$, then $X - M \sim$ exp with mean λ

So $E(X \mid X^{(M)} = M) = M + \lambda$.

Next, if $X^{(M)} < M$ then $X = X^{(M)}$ with probability one

So $E(X \mid X^{(M)}) = X^{(M)}$ if $X^{(M)} < M$.

Thus $h(x) = x$ if $x < M$ and $M + \lambda$ if $x = M$.

5. i) $E Z_n^{(M)} = E(X_1^{(M)}) = \lambda(1 - e^{-M/\lambda})$

So "bias" $= -\lambda e^{-M/\lambda} \to 0$ as $M \to \infty$

ii) $V(Z_n^{(M)}) = \frac{1}{n} V(X_1^{(M)}) \le \frac{1}{n} E(X_1^{(M)})^2 \le \frac{1}{n} 2\lambda^2$

So by Chebychev

$P(|Z_n^{(M)} - E Z_n^{(M)}| > \epsilon) \le \frac{2\lambda^2}{n\epsilon^2} \to 0$

So $Z_n^{(M)} - E Z_n^{(M)} \xrightarrow{p} 0$. But $E Z_n^{(M)} \to \lambda$ by 5i)

So $Z_n^{(M)} - \lambda \xrightarrow{p} 0$.

iii) Since $Z_n^{(M)} - \overline{X}_n = \frac{1}{n}\sum_{j=1}^{n}(X_j^{(M)} - X_j) = Z_n^{(M)} - \lambda + \lambda - \overline{X}_n$

and both $Z_n^{(M)} - \lambda$ and $\overline{X}_n - \lambda \xrightarrow{p} 0$,

and since $\sqrt{n}\frac{(\overline{X}_n - \lambda)}{\lambda} \xrightarrow{d} N(0,1)$ by the Central limit theorem, guess that $\sqrt{n}\frac{(Z_n^{(M)} - \lambda)}{\lambda} \xrightarrow{d} N(0,1)$.