1. (More Estimation in a Zero-In‡ ated Poisson Model) Consider the situation of Problem 1 of Assignment 9, and in particular, inference based on the $n = 20$ observations Vardeman simulated from the distribution.

(a) Find a large sample 90% joint confidence region for $(p, \lambda)$ based on the loglikelihood function (based on inverting LRT’s). Plot this in the $(p, \lambda)$-plane and to the extent possible, compare it to the elliptical region you found in Assignment 9.

(b) Note that for a fixed value of $\lambda$,

$$p = \frac{n_0 - 1}{n \exp(-\lambda) - 1}$$

maximizes the likelihood. Use this fact to find and plot the profile loglikelihood for $\lambda$. Use this plot and make an approximate 90% confidence interval for $\lambda$. How does this interval compare to the one you found in part e) of Problem 1 from Assignment 9?

2. Consider again the model of Problem 2 of Assignment 9. Below are $n = 20$ observations that Vardeman simulated from this model.

$$1.36, 1.35, 0.78, 1.85, 2.32, 0.55, 1.07, -0.57, -0.38, 0.25,$$
$$-0.36, 1.71, 1.40, 0.46, 3.16, -0.78, 0.69, -0.03, 1.26, 0.44$$

(a) Plot the loglikelihood for this sample. What, approximately, is the maximum likelihood estimate for $\alpha$?

(b) If you wished to test the hypothesis $H_0: \alpha = .4$ with Type I error probability $.1$, what would be your decision here? Carefully explain. (Use a likelihood ratio test).

(c) Use the plot from a) and make an approximate 90% confidence interval for $\alpha$ based on the likelihood function (based on inverting LRT’s). Use the method of f) of Problem 2 on Assignment 9 and make another approximate 90% interval. How do these 2 intervals compare?

3. Suppose that $X$, $Y$ and $Z$ are independent binomial variables, $X \sim \text{bin}(n, p_1)$, $Y \sim \text{bin}(n, p_2)$ and $Z \sim \text{bin}(n, p_3)$. For the parameter space (for $(p_1, p_2, p_3)$) $\Theta = [0, 1]^3$, we will consider testing $H_0: p_1 = p_2 = p_3$ based on $(X, Y, Z)$.

(a) Find the general forms of the likelihood ratio tests, the Wald tests and the score tests of this hypothesis.
(b) Use the fact that the parameter space here is basically 3-dimensional while $\Theta_0$ is basically 1-dimensional so that there are 2 independent constraints involved and the limiting $\chi^2$ distributions of the test statistics thus have $\nu = 2$ associated degrees of freedom to actually carry out these tests with $\alpha \approx .05$ if $X = 33, Y = 53$ and $Z = 59$, all based on $n = 100$.

The Following Are Problems About the Large Sample Behavior of Posterior Distributions and Likely Will NOT be Covered on the Stat 543 Final Exam.

1. Problems 1.2.4 and 1.2.5 of B&D.

2. Consider Bayesian inference for the binomial parameter $p$. In particular, for sake of convenience, consider the Uniform $(0, 1)$ (Beta($\alpha, \beta$) for $\alpha = \beta = 1$) prior distribution.

(a) It is possible to argue from reasonably elementary principles that in this binomial context, where $\Theta = (0, 1)$, the Beta posteriors have a consistency property. That is, simple arguments can be used to show that for any fixed $p_0$ and any $\epsilon > 0$, for $X_n \sim \text{binomial} \ (n, p_0)$, the random variable

$$Y_n = \int_{p_0 - \epsilon}^{p_0 + \epsilon} \frac{1}{B(\alpha + X_n, \beta + (n - X_n))} p^{\alpha + X_n - 1}(1 - p)^{\beta + (n - X_n) - 1} dp$$

(which is the posterior probability assigned to the interval $(p_0 - \epsilon, p_0 + \epsilon)$) converges in $p_0$ probability to 1 as $n \to \infty$. This part of the problem is meant to lead you through this argument. Let $\epsilon > 0$ and $\delta > 0$.

i) Argue that there exists $m$ such that if $n \geq m$, \[ \left| \frac{x_n}{n} - \frac{\alpha + x_n}{\alpha + \beta + n} \right| < \frac{\epsilon}{3} \ \forall x_n = 0, 1, \ldots, n. \]

ii) Note that the posterior variance is \[ \frac{(\alpha + x_n)(\beta + n - x_n)}{(\alpha + \beta + n)^2 (\alpha + \beta + n + 1)}. \] Argue there is an $m'$ such that if $n \geq m'$ the probability that the posterior assigns to \[ \left( \frac{\alpha + x_n}{\alpha + \beta + n} - \frac{\epsilon}{3}, \frac{\alpha + x_n}{\alpha + \beta + n} + \frac{\epsilon}{3} \right) \]

is at least $1 - \delta$ $\forall x_n = 0, 1, \ldots, n$.\n
iii) Argue that there is an $m''$ such that if $n \geq m''$ the $p_0$ probability that \[ \left| \frac{x_n}{n} - p_0 \right| < \frac{\epsilon}{3} \] is at least $1 - \delta$.

Then note that if $n \geq \max(m, m', m'')$ i) and ii) together imply that the posterior probability assigned to \( \left( \frac{x_n}{n} - \frac{2\epsilon}{3}, \frac{x_n}{n} + \frac{2\epsilon}{3} \right) \) is at least $1 - \delta$ for any realization $x_n$. Then provided $\left| \frac{x_n}{n} - p_0 \right| < \frac{\epsilon}{3}$ the posterior probability assigned to $(p_0 - \epsilon, p_0 + \epsilon)$ is also at least $1 - \delta$. But iii) says this happens with $p_0$ probability at least $1 - \delta$. That is, for large $n$, with $p_0$ probability at least $1 - \delta$, $Y_n \geq 1 - \delta$. Since $\delta$ is arbitrary, (and $Y_n \leq 1$) we have the convergence of $Y_n$ to 1 in $p_0$ probability.

(b) Vardeman intends to argue in class that posterior densities for large $n$ tend to look normal (with means and variances related to the likelihood material). The posteriors in this binomial problem are Beta $(\alpha + x_n, \beta + (n - x_n))$ (and we can think of $X_n \sim \text{Bi} \ (n, p_0)$ as derived as the sum of $n$ iid Bernoulli $(p_0)$ variables). So we ought to expect Beta distributions for large parameter values to look roughly normal. To illustrate this do the following. For $\rho = .3$ (for example ... any
other value would do as well), consider the Beta \((\alpha + n\rho, \beta + n(1 - \rho))\) (posterior) distributions for \(n = 10, 20, 40\) and \(100\). For \(p_n \sim \text{Beta}(\alpha + n\rho, \beta + n(1 - \rho))\) plot the probability densities for the variables

\[
\sqrt{\frac{n}{\rho(1 - \rho)}} (p_n - \rho)
\]
on a single set of axes along with the standard normal density. Note that if \(W\) has pdf \(f(\cdot)\), then \(aW + b\) has pdf \(g(\cdot) = \frac{1}{a} f\left(\frac{a - b}{a}\right)\). (Your plots are translated and rescaled posterior densities of \(p\) based on possible observed values \(x_n = .3n\).)

If this is any help in doing this plotting, Vardeman tried to calculate values of the Beta function using MathCad and got the following: \((B(4, 8))^{-1} = 1.32 \times 10^3\), \((B(7, 15))^{-1} = 8.14 \times 10^5\), \((B(13, 29))^{-1} = 2.291 \times 10^{11}\) and \((B(31, 71))^{-1} = 2.967 \times 10^{27}\).