HW 2

1] Find the Maximum Likelihood Estimator for the following:

a) The probability of success p in the Bernoulli(p) model

* To find an MLE, first identify the model's pmf

Bernoulli $\longrightarrow f(k, p) = \begin{cases} p & \text{if } k=1 \\ 1-p & \text{if } k=0 \end{cases}$

or, since a Bernoulli distribution is a Binomial one with $n=1$:

$f(k,p) = p^k (1-p)^{1-k}$ for $k \in \{0,1\}$

* The likelihood function is the total probability of "finding such a sample." We identify the likelihood function by taking the cumulative product of the probability of all the outcomes. At this point, we evaluate the likelihood function using only parameter p, as we now consider k a "constant" of sorts, as we are evaluating the likelihood of p representing this particular "sample."

$L(p) = \prod p^k (1-p)^{1-k}$

* We use the log-likelihood function to make computations easier

$\ln L(p) = \ln(\prod_{k=1} p^k (1-p)^{1-k}) = \sum \ln(p^k (1-p)^{1-k}) = \sum [\ln(p^k) + \ln((1-p)^{1-k})]$

$= \sum_{i=0} [k \ln(p) + (1-k)\ln(1-p)] = S_k \ln(p) + (n-S_k)\ln(1-p)$

* Now that we have a simple log-likelihood function, we differentiate it with respect to p and set the result equal to zero to find the maximum:

$\frac{d \ln L}{dp} = S_k \cdot \frac{1}{p} - (n-S_k) \cdot \frac{1}{1-p} = \frac{S_k}{p} - \frac{n-S_k}{(1-p)} = \frac{S_k(1-p) - p(1-S_k)}{p(1-p)}$

$= \frac{S_k - pS_k - np + pS_k}{p(1-p)} = \frac{S_k - pn}{p(1-p)}$

* Set the expression equal to zero and solve for p to find extreme values of p

$0 = \frac{S_k - np}{p(1-p)} \longrightarrow \hat{p} = \frac{S_k}{n}$

* Differentiate again and set to zero to determine if it is a max

$\frac{\partial}{\partial p}$

* The above result will always be negative, the estimator is a maximum.

b) The probability of successes p in a Binomial(n,p) model:

* The pmf

$$f(n,p) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

* The likelihood function

$$L(p|n) = \prod_{i=1}^{n} \frac{n!}{k_i(n-k_i)} p^{k_i}(1-p)^{n-k_i}$$

* The log-likelihood function

$$\ln L(p|n) = \ln\left(\prod_{i=1}^{n} \frac{n!}{k_i!(n-k_i)} p^{k_i}(1-p)^{n-k_i}\right) = \sum_{i=1}^{n} \ln\left(\frac{n!}{k_i!(n-k_i)!} p^{k_i}(1-p)^{n-k_i}\right)$$

$$= \sum_{i=1}^{n} \ln\binom{n}{k_i} + \ln(p^{k_i}) + \ln\left((1-p)^{n-k_i}\right) = \sum_{i=1}^{n} \ln\binom{n}{k} + k_i \ln(p) + (n-k_i)\ln(1-p)$$

* Differentiate the expression with respect to our target parameter:

$$\frac{d \ln L(p)}{d p} = \sum 0 + k \cdot \frac{1}{p} - (n-k) \cdot \frac{1}{1-p} = \frac{k}{p} - \frac{(n-k)}{(1-p)} = \frac{k(1-p)-(n-k)p}{p(1-p)}$$

*Note: We can drop the Σ because I am now asuming k represents $\sum_{i=1}^{n} k$

* Find an expression for an extreme value

$$0 = k(1-p)-(n-k)p^{\dagger} \rightarrow (n-k)p = k(1-p) \rightarrow pn - pk = k - pk \rightarrow pn = k$$

$p = \frac{k}{n}$, which is how we define $\hat{p}$ $\rightarrow p = \frac{k}{n} = \hat{p}$

* Differentiate once more to ensure a maximum has been found

$$\frac{\partial^2 \ln L(p)}{\partial p^2} = \left[\frac{k(1-p)-(n-k)p}{p(1-p)}\right]' = \frac{-3p^2 n - 2pk}{p^2(p^2-2p+1)}$$

The result of the above expression will always be negative, thus we have found a maximum.

Therefore, the MLE of the Binomial distribution is $\hat{p} = \frac{k}{n}$

† Note: In order for any fraction to equal zero, it's numerator must also be equal to zero, so for simplicity I dropped the $p(1-p)$ denominator.

c) The probability of success p in the Geometric model

* The pmf:

$$f(k, p) = p(1-p)^{k-1}$$

* The likelihood → the log-likelihood

$$L(p) = \prod_{i=1}^{n} p(1-p)^{k_i - 1} \longrightarrow \ln L(p) = \sum_{i=1}^{n} \ln\left(p(1-p)^{k_i-1}\right)$$

$$= \sum\left[\ln(p) + \ln\left((1-p)^{k_i-1}\right)\right] = \sum_{i=1}^{n}\left[\ln(p) + (k-1)\ln(1-p)\right]$$

$$= n \cdot \ln(p) + (S_k - n)\ln(1-p)$$

* Differentiate with respect to p

$$\frac{d \ln L(p)}{dp} = n \cdot \frac{1}{p} - S_k - n \cdot \frac{1}{1-p} = \frac{n}{p} - \frac{S_k - n}{1-p} = \frac{n(1-p) - p(S_k - n)}{p(1-p)}$$

$$\rightarrow n - np - S_k p + np = n - S_k p^{+}$$

* Evaluate for extreme

$$0 = n - S_k p \rightarrow S_k p = n \rightarrow p = \frac{n}{S_k}$$

* Ensure the extreme is a max

$$\frac{d^2 \ln L(p)}{dp^2} = \frac{2pn - n}{p^2(1-p)}$$ which will always be negative → we have a max

Thus, our estimator, $\hat{p}$, is the result of the number of samples collected, divided by the total number of Bernoulli trials required for a success in each sample, or the total number of successes observed (1 for each trial) divided by the total number of trials (not Bernoulli ones, but each series of Bernoulli trials conducted until a success is seen).

$^{+}$ Again, I dropped the denominator in preparation for extreme value finding

d) The intensity λ in the Poisson (λ) model.

* pmf

$$f(\lambda, k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

* Likelihood function

$$L(\lambda) = \prod_{i=1}^{n} e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

* Log-likelihood function

$$\ln L(\lambda) = \ln\left(\prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{k_i}}{k_i!}\right) = \sum_{i=1}^{n} \ln\left(e^{-\lambda} \frac{\lambda^{k_i}}{k_i!}\right) = \sum_{i=1}^{n} \left[\ln(e^{-\lambda}) + \ln\left(\frac{\lambda^{k_i}}{k_i!}\right)\right]$$

$$= \sum \left[-\lambda \ln(e) + \ln(\lambda^{k_i}) - \ln(k_i!)\right] = \sum_{i=1}^{n} \left[k_i \ln(\lambda) - \ln(k_i!) - \lambda \right]$$

* Differentiate and find an extreme

$$\frac{d \ln L(\lambda)}{d\lambda} = \sum_{i=1}^{n} k_i \cdot \frac{1}{\lambda} - 0 - 1 = \frac{S_k}{\lambda} - n$$

$$0 = \frac{S_k}{\lambda} - n \rightarrow \lambda = \frac{S_k}{n} \longrightarrow \hat{\lambda} = \frac{S_k}{n}$$

* Differentiate again to ensure max

$$\frac{d^2 \ln L(\lambda)}{d\lambda^2} = -S_k \lambda^{-2},$$ This expression will always be negative because λ^2 is always positive, and S_k will always be positive

Thus we have our estimator $\hat{\lambda}$ which is found by the ratio of the total number of occurrences of interest found in all samples/trials and the total number of samples/trials.

2] Find the Fisher information for the following:

a) The Bernoulli model:

*For simplicity, let $F = n - k$ (or $F = 1 - k$)

$$I(p) = -E\left[\frac{\partial^2}{\partial p^2} \ln(f(k;p)) \,\Big|\, p\right] = -E\left[\frac{\partial^2}{\partial p^2} \ln\left(\binom{n}{k} p^k (1-p)^F\right) \Big|\, p\right]$$

$$= -E\left[\frac{\partial^2}{\partial p^2} \ln\binom{n}{k} + \ln(p^k) + \ln((1-p)^F) \,\Big|\, p\right] = -E\left[\frac{\partial^2}{\partial p^2} \ln\binom{n}{k} + k\ln(p) + F\ln((1-p)) \,\Big|\, p\right]$$

$$= -E\left[\frac{\partial}{\partial p} \frac{k}{p} - \frac{F}{(1-p)} \,\Big|\, p\right] = E\left[-\frac{k}{p^2} + \frac{F}{(1-p)^2} \,\Big|\, p\right]$$

*Now, because p is a condition of this expression, we can evaluate for expected values in the expression, specifically $E[k]$ and $E[F]$. We can use the accepted values of a Bernoulli model because p is a condition, i.e. $-E[k] = -np$ and $E[F] = -(1-p)$. Due to the linear properties of expected values, this should not otherwise alter the expression.

$$= \frac{np}{p^2} + \frac{np(1-p)}{(1-p)^2} = \frac{np(1-p)^2 + np^2(1-p)}{p^2(1-p)^2}$$

* For a Bernoulli model $n=1$, so I will make this substitution now:

$$= \frac{(1)p(1-p)^2 + (1)p^2(1-p)}{p^2(1-p)^2} = \frac{p(1-p)^2 + p^2(1-p)}{p^2(1-p)^2} = \frac{p(1-2p+p^2) + p^2 - p^3}{p \cdot p(1-p)(1-p)}$$

$$= \frac{p - 2p^2 + p^3 + p^2 - p^3}{p(1-p)(p-p^2)} = \frac{p - p^2}{p(1-p)(p-p^2)} = \frac{1}{p(1-p)}$$

Thus, $I(p) = \dfrac{1}{p(1-p)}$

b) The probability of success p in the Binomial model

* This is the same as the Bernoulli model, except n is a variable and not equal to 1

$$I(p) = -E\left[\frac{\partial^2}{\partial p^2} \ln\left(f(k;p)\right) \Big| p\right] = -E\left[\frac{\partial^2}{\partial p^2} \ln\left(\binom{n}{k} p^k (1-p)^{n-k}\right) \Big| p\right]$$

$$= -E\left[\frac{\partial^2}{\partial p^2} \ln\binom{n}{k} + k \ln(p) + (n-k)\ln((1-p)) \Big| p\right] = -E\left[\frac{\partial}{\partial p} \frac{k}{p} - \frac{n-k}{1-p} \Big| p\right]$$

$$= E\left[-\frac{k}{p^2} - \frac{n-k}{(1-p)^2} \Big| p\right]$$

* Now substitute $E[k] = np$ and $E[n-k] = n(1-p)$

$$= -\left[-\frac{np}{p^2} - \frac{n(1-p)}{(1-p)^2}\right] = \left[\frac{n}{p} + \frac{n}{(1-p)}\right] = -\left[\frac{n - np + np}{p(1-p)}\right]$$

$$= \left[\frac{n}{p(1-p)}\right]$$

Thus $I(p) = \dfrac{n}{p(1-p)}$

c) The probability of success p in the Geometric model.

$$I(p) = E\left[\left(\frac{\partial}{\partial p}\ln(f(k;p))\right)^2 \middle| p\right] = E\left[\left(\frac{\partial}{\partial p}\ln\left(p(1-p)^{k-1}\right)\right)^2 \middle| p\right]$$

$$= E\left[\left(\frac{\partial}{\partial p}\ln(p) + (k-1)\ln((1-p))\right)^2 \middle| p\right] = E\left[\left(\frac{1}{p} - \frac{k-1}{1-p}\right)^2 \middle| p\right]$$

$$= E\left[\frac{1}{p^2} - 2\left(\frac{k-1}{p(1-p)}\right) + \frac{(k-1)^2}{(1-p)^2} \middle| p\right]$$

* Replace $E[k-1]$ with $\frac{1}{p}$ and solve for p

$$= \frac{1}{p^2} - 2\left(\frac{1}{p\,p(1-p)}\right) + \left(\frac{1}{p}\right)^2\left(\frac{1}{(1-p)^2}\right) = \frac{1}{p^2} - \frac{2}{p^2+2p^3+p^4} + \frac{1}{p+2p^2+p^3}$$

$$= \frac{1}{(1-p)^2} \quad ?$$

$$I(p) = \frac{p^2}{(1-p)}$$

d) The Intensity λ in the Poisson model

$$I(\lambda) = -E\left[\frac{\partial^2}{\partial\lambda^2} \ln(f(k;\lambda)) \middle| \lambda\right] = -E\left[\frac{\partial^2}{\partial\lambda^2} \ln\left(\frac{\lambda^k}{k!} \cdot e^{-\lambda}\right) \middle| \lambda\right]$$

$$= -E\left[\frac{\partial^2}{\partial\lambda^2} k(\ln(\lambda) - \ln(k!)) + -\lambda \ln(e) \middle| \lambda\right] = -E\left[\frac{\partial^2}{\partial\lambda^2} k \ln(\lambda) - k\ln(k!) - \lambda(1) \middle| \lambda\right]$$

$$= -E\left[\frac{\partial}{\partial\lambda} \frac{k}{\lambda} - 0 - 1 \middle| \lambda\right] = -E\left[-\frac{k}{\lambda^2} \middle| \lambda\right]$$

* Substitute: $E[k] = \lambda$

$$= -\left(-\frac{(\lambda)}{\lambda^2}\right) = \frac{1}{\lambda}$$

Thus $I(\lambda) = \frac{1}{\lambda}$

3| Assess the efficiency of the MLE estimators from #1

$$e(\hat{\theta}) = \frac{I(\theta)^{-1}}{var(\hat{\theta})}$$

a) The Bernoulli model:

I found for $I(p)$: $\frac{1}{p(1-p)}$ MLE(p): $\hat{p} = \frac{S_k}{n} \rightarrow var(\hat{p}) = \hat{p}(1-\hat{p})$

$$e(\hat{p}) = \frac{\left(\frac{1}{p(1-p)}\right)^{-1}}{\hat{p}(1-\hat{p})} \approx 1$$

b) The Binomial model:

$I(p) = \frac{n}{p(1-p)}$ MLE$(p) = \hat{p} = \frac{k}{n} \rightarrow var(\hat{p}) = n\hat{p}(1-\hat{p})$

$$e(\hat{p}) = \frac{\frac{p(1-p)}{n}}{n\hat{p}(1-\hat{p})} \approx \frac{n}{n} = 1$$

c) The Geometric model

$I(p)$: $\frac{p^2}{1-p}$

MLE: $\hat{p} = \frac{n}{S_k} \rightarrow var(\hat{p}) = \frac{1-\hat{p}}{\hat{p}^2}$

$$e(\hat{p}) = \frac{\left(\frac{p^2}{1-p}\right)^{-1}}{\left(\frac{1-\hat{p}}{\hat{p}^2}\right)} \approx 1$$

d) The Poisson model:

$I(\lambda)$: $\frac{1}{\lambda}$

MLE: $\hat{\lambda} = \frac{k}{n} \rightarrow var(\hat{\lambda}) = \hat{\lambda}$

$$e(\hat{\lambda}) = \frac{\left(\frac{1}{\lambda}\right)^{-1}}{(\hat{\lambda})} = \frac{\lambda}{\hat{\lambda}} \approx 1$$