ECO 520 - Fall 2020
Final Exam

Haixiang Zhu
SBU ID#113029589

December 7, 2020

1. (a) *Proof.*

$$\because \frac{1}{n}\sum_{i=1}^{n} w_i = 1, E(X) = \mu, Var(X) = \sigma^2$$

$$\begin{aligned}
\therefore E[\bar{X}_n^*] &= E\left[\frac{1}{n}\sum_{i=1}^{n} w_i X_i\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} w_i E[X_i] \\
&= \mu
\end{aligned}$$

□

(b)

$$\begin{aligned}
Var(\bar{X}_n^*) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} w_i X_i\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} w_i^2 Var(X_i) \\
&= \frac{\sigma^2}{n^2}\sum_{i=1}^{n} w_i^2
\end{aligned}$$

(c) *Proof.* Applying Chebyshev's inequlity

$$
\begin{aligned}
&P(|\bar{X}_n^* - \mu| \geq \varepsilon) \leq \frac{\sigma^2 \sum_{i=1}^{n} w_i^2}{n^2 \varepsilon^2} \qquad\qquad \forall \varepsilon > 0 \\
&\lim_{n\to\infty} \frac{\sigma^2}{n^2} \sum_{i=1}^{n} w_i^2 = 0 \\
\Rightarrow &P(|\bar{X}_n^* - \mu| \geq \varepsilon) = 0 \\
\Rightarrow &\bar{X}_n^* \xrightarrow{p} \mu
\end{aligned}
$$

□

The additional condition for w_i is

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} w_i^2 = 0$$

(d) Directly using ML theory

$$\sqrt{n}(\bar{X}_n^* - \mu) \xrightarrow{d} N\left(0, \sigma^2 \sum_{i=1}^{n} w_i^2\right)$$

2. By definition of pdf

$$\int_0^\infty \frac{1}{\sqrt{2\pi}\Phi(\lambda)} e^{-\frac{(x-\lambda)^2}{2}} \mathrm{d}x = 1$$

(a) MGF

$$
\begin{aligned}
M_X(t) &= E[e^{tX}] \\
&= \int_{-\infty}^{\infty} e^{tx} f_X(x) \mathrm{d}x \\
&= \int_0^\infty \frac{e^{tx}}{\sqrt{2\pi}\Phi(\lambda)} e^{-\frac{(x-\lambda)^2}{2}} \mathrm{d}x \\
&= \frac{\Phi(\lambda+t)}{\Phi(\lambda)} e^{\lambda t + \frac{t^2}{2}} \int_0^\infty \frac{1}{\sqrt{2\pi}\Phi(\lambda+t)} e^{-\frac{[x-(\lambda+t)]^2}{2}} \mathrm{d}x \\
&= \frac{\Phi(\lambda+t)}{\Phi(\lambda)} e^{\lambda t + \frac{t^2}{2}} \qquad (1)
\end{aligned}
$$

(b)

$$
\begin{aligned}
E[X] &= M_X'(0) \\
&= \frac{1}{\Phi(\lambda)}[\phi(\lambda+t) + (\lambda+t)\Phi(\lambda+t)]e^{\lambda t + \frac{t^2}{2}}\Big|_{t=0} \\
&= \underbrace{\frac{1}{\sqrt{2\pi}\Phi(\lambda)}e^{-\frac{\lambda^2}{2}}}_{A} + (\lambda+t)M_X(t)\Bigg|_{t=0} \\
&= \frac{1}{\sqrt{2\pi}\Phi(\lambda)}e^{-\frac{\lambda^2}{2}} + \lambda \\
E[X^2] &= M_X''(0) \\
&= M_X(t) + (\lambda+t)M_X'(t)|_{t=0} \\
&= 1 + \lambda(\lambda + A) \\
&= 1 + \lambda^2 + \frac{\lambda}{\sqrt{2\pi}\Phi(\lambda)}e^{-\frac{\lambda^2}{2}} \\
E[X^3] &= M_X'''(0) \\
&= 2M_X'(t) + (\lambda+t)M_X''(t)|_{t=0} \\
&= 2(\lambda + A) + \lambda[1 + \lambda(\lambda + A)] \\
&= 3\lambda + \lambda^3 + \frac{2+\lambda^2}{\sqrt{2\pi}\Phi(\lambda)}e^{-\frac{\lambda^2}{2}}
\end{aligned}
$$

(c) CF

$$
\begin{aligned}
\phi_X(u) &= E[e^{iuX}] \\
&= \int_{-\infty}^{\infty} e^{iux} f_X(x)\mathrm{d}x \\
&\xlongequal{\text{plug } t=iu \text{ into (1)}} \frac{\Phi(\lambda+iu)}{\Phi(\lambda)}e^{\lambda iu - \frac{u^2}{2}}
\end{aligned}
$$

3. Negative binomial distribution
Note that

$$
\begin{aligned}
\binom{m+x-1}{m-1} &= \binom{m+x-1}{x} = \frac{(m+x-1)(m+x-2)\cdots(m+1)m}{x!} \\
&= (-1)^x \frac{(-m)(-m-1)\cdots(-m-x+1)}{x!} \\
&= (-1)^x \binom{-m}{x} \\
\Rightarrow \binom{-m}{x} &= (-1)^x \binom{m+x-1}{x}
\end{aligned}
$$

Therefore

$$
\begin{aligned}
(1+t)^{-m} &= \sum_{x=0}^{\infty} \binom{-m}{x} t^x \\
&= \sum_{x=0}^{\infty} \binom{m+x-1}{x} (-t)^x \qquad (2)
\end{aligned}
$$

(a) MGF

$$
\begin{aligned}
M_X(t) =& E[e^{tX}] \\
=& \sum_x e^{tx} P(X=x) \\
=& \sum_{x=0}^{\infty} \binom{m+x-1}{x} p^m (1-p)^x e^{tx} \\
\xlongequal{\text{plugging (2)}}& p^m [1-(1-p)e^t]^{-m} \\
=& \left[\frac{p}{1-(1-p)e^t}\right]^m \qquad (3)
\end{aligned}
$$

(b)

$$
\begin{aligned}
E[X] &= M'_X(0) \\
&= m(1-p)e^t p^m [1-(1-p)e^t]^{-m-1}\big|_{t=0} \\
&= \frac{1-p}{p} m \\
E[X^2] &= M''_X(0) \\
&= m(1-p)p^m \{e^t [1-(1-p)e^t]^{-m-1} + e^{2t}(1+m)(1-p)[1-(1-p)e^t]^{-m-2})\}\big|_{t=0} \\
&= m(1-p)[p^{-1} + (1+m)(1-p)p^{-2}] \\
&= \frac{m(1-p)(1+m-mp)}{p^2} \\
E[X^3] &= M'''_X(0) \\
&= m(1-p)p^m \{e^t [1-(1-p)e^t]^{-m-1} + 3e^{2t}(1+m)(1-p)[1-(1-p)e^t]^{-m-2}) \\
&\qquad + e^{3t}(1+m)(2+m)(1-p)^2 [1-(1-p)e^t]^{-m-3}\}\big|_{t=0} \\
&= m(1-p)[p^{-1} + 3(1+m)(1-p)p^{-2} + (1+m)(2+m)(1-p)^2 p^{-3}] \\
&= \frac{m(1-p)(1+m-mp)(2+m-mp)}{p^3} - \frac{m(1-p)}{p^2}
\end{aligned}
$$

(c) CF

$$\begin{aligned}
\phi_X(u) &= E[e^{iuX}] \\
&= \sum_x e^{iux} P(X = x) \\
&\xlongequal{\text{plug } t=iu \text{ into (3)}} \left[\frac{p}{1-(1-p)e^{iu}}\right]^m
\end{aligned}$$

4. Geometric distribution

(a) The likelihood and log-likelihood are

$$\begin{aligned}
L_X(p) &= \prod_{i=1}^{n} p^{x_i}(1-p) \\
l_X(p) &= \log(p)\sum_{i=1}^{n} x_i + n\log(1-p)
\end{aligned}$$

FOC

$$\frac{1}{p}\sum_{i=1}^{n} x_i - \frac{n}{1-p} = 0$$

Hence, the MLE is

$$\hat{p}_n = \frac{\bar{x}}{\bar{x}+1}$$

where $\bar{x} = \frac{1}{n}\sum\limits_{i=1}^{n} x_i$.

The second derivative is given by

$$\begin{aligned}
H_n(p) &= \frac{\partial^2 l_X(p)}{\partial p^2} \\
&= -n\left[\frac{\bar{x}}{p^2} + \frac{1}{(1-p)^2}\right] < 0 \\
H_n(\hat{p}_n) &= -n\left[\frac{1}{\hat{p}_n(1-\hat{p}_n)} + \frac{1}{(1-\hat{p}_n)^2}\right] \\
&= -\frac{n}{\hat{p}_n(1-\hat{p}_n)^2} < 0
\end{aligned}$$

which turns out that the MLE $\hat{p}_n$ is the unique global maxima.

(b) *Proof.* Since the geometric distribution is a special case of the negative binomial distribution with $m = 1$, the MGF is

$$M_X(t) = \frac{1-p}{1-pe^t}$$

Then

$$\begin{cases} E(X) = M'_X(0) = \dfrac{p}{1-p} < \infty \\ E(X^2) = M''_X(0) = \dfrac{p(1+p)}{(1-p)^2} < \infty \end{cases}$$

Applying WLLN

$$\bar{x} \xrightarrow{p} E(x) = \frac{p}{1-p}$$

$$\Rightarrow \hat{p}_n = \frac{\bar{x}}{\bar{x}+1} \xrightarrow{p} \frac{\frac{p}{1-p}}{\frac{p}{1-p}+1} = p$$

$\square$

(c) The information matrix is

$$I_n(\hat{\rho}_n) = -E[H_n(\hat{p}_n)] = \frac{1}{\hat{p}_n(1-\hat{p}_n)^2}$$

Directly using ML theory

$$\sqrt{n}(\hat{p}_n - p) \xrightarrow{d} N\left(0, p(1-p)^2\right)$$

5. Zero-inflated Poisson distribution

(a) Define n_0 to be the number of X_i's taking the value 0. Since this is a zero-modified distribution, one of the ML equations is

$$p + (1-p)e^{-\lambda} = \frac{n_0}{n} \tag{4}$$

where $\dfrac{n_0}{n}$ is the observed proportion of zeros.
The log-likelihood function is (ignoring any constant term)

$$l_X(p,\lambda) = n_0\{\log[p+(1-p)e^{-\lambda}]\} + (n-n_0)[\log(1-\pi)-\lambda] + n\bar{x}\log(\lambda)$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.
FOC

$$\begin{cases} \dfrac{\partial l_X(p,\lambda)}{\partial p} = n_0 \dfrac{1-e^{\lambda}}{p+(1-p)e^{-\lambda}} - \dfrac{n-n_0}{1-p} = 0 \\ \dfrac{\partial l_X(p,\lambda)}{\partial \lambda} = -n_0 \dfrac{(1-p)e^{\lambda}}{p+(1-p)e^{-\lambda}} - (n-n_0) + \dfrac{n\bar{x}}{\lambda} = 0 \end{cases}$$

Along with (4)

$$1 - \hat{p}_n = \frac{\bar{x}}{\hat{\lambda}_n}$$

Substituting out $\hat{p}_n$ gives

$$\bar{x}(1-e^{-\hat{\lambda}_n}) = \hat{\lambda}_n \left(1 - \frac{n_0}{n}\right)$$

and hence $\hat{\lambda}_n$ (and $\hat{p}_n$) can be obtained by iteration.
The second derivative is given by

$$\begin{cases} H_n(p) = \dfrac{\partial^2 l_X(p,\lambda)}{\partial p^2} = -\dfrac{n-n_0}{(1-p)^2} < 0 \\ H_n(\lambda) = \dfrac{\partial^2 l_X(p,\lambda)}{\partial \lambda^2} = n(1-p)e^{-\lambda} - \dfrac{n\bar{x}}{\lambda^2} \text{ (undetermined)} \end{cases}$$

which turns out that MLE $\hat{p}_n$ is the unique global maxima. So is MLE $\hat{\lambda}_n$ because of the one-to-one mapping.

(b) The MLEs are consistent, intuitively, which guides us to determine the asymptotic distribution.

(c) The information matrices are

$$\begin{cases} I_n(\hat{p}_n) = -E[H_n(\hat{p}_n)] = \dfrac{1-\frac{n_0}{n}}{(1-\hat{p}_n)^2} = \dfrac{1}{(1-\hat{p}_n)\left(1-e^{-\hat{\lambda}_n}\right)} \\ I_n(\hat{\lambda}_n) = -E[H_n(\hat{\lambda}_n)] = -(1-\hat{p}_n)e^{-\hat{\lambda}_n} + \dfrac{\bar{x}}{\hat{\lambda}_n^2} = \dfrac{(1-\hat{p}_n)\left(1-\hat{\lambda}_n e^{-\hat{\lambda}_n}\right)}{\hat{\lambda}_n} \end{cases}$$

Directly using ML theory

$$\begin{cases} \sqrt{n}(\hat{p}_n - p) \xrightarrow{d} N\left(0, (1-p)\left(1-e^{-\lambda}\right)\right) \\ \sqrt{n}(\hat{\lambda}_n - \lambda) \xrightarrow{d} N\left(0, \dfrac{\lambda}{(1-p)\left(1-\lambda e^{-\lambda}\right)}\right) \end{cases}$$