SKETCH of the SOLUTIONS
STK4051 - spring 2025

Exercise 1

a Using the moment generating function,

$$\begin{aligned} E[Y] &= \left.\frac{d\, M_Y(t)}{d\, t}\right|_{t=0} \\ &= \left.\frac{d\left(1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}\right)}{d\, t}\right|_{t=0} \\ &= \left.\sum_{k=1}^{\infty} k\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^{k-1}}{k!}\right|_{t=0}. \end{aligned}$$

Since all terms t^k are 0 but when $k=1$, it becomes

$$E[Y] = \left(\prod_{r=0}^{0}\frac{\alpha+r}{\alpha+\beta+r}\right) = \frac{\alpha}{\alpha+\beta}.$$

For the variance,

$$\begin{aligned} E[Y^2] &= \left.\frac{d^2\, M_Y(t)}{d\, t^2}\right|_{t=0} \\ &= \left.\frac{d\,\sum_{k=1}^{\infty} k\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^{k-1}}{k!}}{d\, t}\right|_{t=0} \\ &= \left.\sum_{k=1}^{\infty} k(k-1)\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^{k-2}}{k!}\right|_{t=0} \end{aligned}$$

where now all terms of the sum are 0 but for $k=2$, so,

$$\begin{aligned} E[Y^2] &= 2\left(\prod_{r=0}^{1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{1}{2!} \\ &= \frac{\alpha}{\alpha+\beta}\frac{\alpha+1}{\alpha+\beta+1} \end{aligned}$$

and

$$\begin{aligned} Var[Y] &= E[Y^2] - (E[Y])^2 \\ &= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 \\ &= \frac{(\alpha^2+\alpha)(\alpha+\beta) - \alpha^2(\alpha+\beta+1)}{(\alpha+\beta)^2(\alpha+\beta+1)} \\ &= \frac{\alpha^3+\alpha^2\beta+\alpha^2+\alpha\beta-\alpha^3-\alpha^2\beta-\alpha^2}{(\alpha+\beta)^2(\alpha+\beta+1)} \\ &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \end{aligned}$$

To compute the moment estimators,

$$\begin{aligned} \bar{y} &= \frac{\alpha}{\alpha+\beta} \\ s^2 &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \end{aligned}$$

From the first formula

$$\alpha+\beta = \frac{\alpha}{\bar{y}}$$

and

$$\beta = \alpha\frac{1-\bar{y}}{\bar{y}}. \tag{1}$$

Substituting in the second formula,

$$\begin{aligned} s^2 &= \frac{\alpha^2\frac{1-\bar{y}}{\bar{y}}}{\frac{\alpha^2}{\bar{y}^2}(\frac{\alpha}{\bar{y}}+1)} \\ &= \frac{\bar{y}^2\frac{1-\bar{y}}{\bar{y}}}{(\frac{\alpha}{\bar{y}}+1)} \end{aligned}$$

and

$$\left(\frac{\alpha}{\bar{y}}+1\right)s^2 = \bar{y}(1-\bar{y})$$

so that

$$\frac{\alpha s^2}{\bar{y}} = \bar{y}(1-\bar{y}) - s^2$$

and

$$\hat{\alpha} = \bar{y}\left(\frac{\bar{y}(1-\bar{y})}{s^2} - 1\right)$$

Substituting in formula (1),

$$\hat{\beta} = \bar{y}\left(\frac{\bar{y}(1-\bar{y})}{s^2} - 1\right)\frac{(1-\bar{y})}{\bar{y}} = \left(\frac{\bar{y}(1-\bar{y})}{s^2} - 1\right)(1-\bar{y}).$$

b One possible way to show that $\log y$ and $\log(1-y)$ are sufficient statistics for α and β, respectively, is to rewrite the density in the form of an exponential family,

$$f(y;\alpha,\beta) = \underbrace{\frac{1-y}{y} I(0<y<1)}_{h(y)} \exp\left\{ \underbrace{\alpha}_{w_1(\alpha)} \underbrace{\log y}_{t_1(y)} + \underbrace{\beta}_{w_2(\beta)} \underbrace{\log(1-y)}_{t_2(y)} - \underbrace{\log \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}}_{\kappa(\alpha,\beta)} \right\}$$

and invoke theorem 6.2.10.

Similarly, when $\alpha = \beta$, rewriting the density as

$$f(y;\alpha,\beta) = \underbrace{\frac{1-y}{y} I(0<y<1)}_{h(y)} \exp\left\{ \underbrace{\alpha}_{w(\alpha)} \underbrace{\log(y(1-y))}_{t(y)} - \underbrace{\log \frac{\Gamma(\alpha)^2}{\Gamma(2\alpha)}}_{\kappa(\alpha)} \right\},$$

shows that the sufficient statistic for the only parameter is $\log(y(1-y))$.

c We first show that $X_{(1)} \sim \text{Beta}(1,n)$ when $X_1,\ldots,X_n$ are independent and identically distributed Beta$(1,1)$ random variables, i.e., Unif$(0,1)$ ($f(y;1,1) = I(0<y<1)$) by using Theorem 5.4.4,

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x)[F_X(x)]^{(j-1)}[1-F_X(x)]^{n-1}$$

which, because $f_X(x) = I(0<x<1)$ and $F_X(x) = x$, gives, with $\alpha = j = 1$ and $\beta = n$, the density of a Beta$(1,n)$. Now

$$\begin{aligned} Pr(nX_{(1)} \le x) &= Pr(X_{(1)} \le x/n) \\ &= Pr(X_{(1)} > x/n) \\ &= 1 - Pr(\prod_{i=1}^{n} X_i > x/n) \end{aligned}$$

where all X_i are i.i.d. Unif$(0,1)$, so

$$Pr(nX_{(1)} \le x) = 1 - (1-x/n)^n$$

and, using the fact that $\lim_{n\to\infty}(1-x/n)^n = e^{-x}$,

$$Pr(nX_{(1)} \le x) = 1 - e^{-x}$$

which is the cumulative distribution function of an exponential with mean 1.

d We simply need to apply Theorem 2.1.5,

$$f(x;\alpha,\beta,a,b) = f(g^{-1}(y))\left|\frac{d\,g^{-1}(y)}{d\,y}\right|$$

where

$$g^{-1}(y) = \frac{y-a}{b-a}$$

and

$$\frac{d\,(y-a)/(b-a)}{d\,y} = \frac{1}{b-a}.$$

Therefore,

$$f(x;\alpha,\beta,a,b) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\left(\frac{y-a}{b-a}\right)^{\alpha-1}\left(1-\frac{y-a}{b-a}\right)^{\beta-1}\frac{1}{b-a} & \text{if } a<y<b \\ 0 & \text{otherwise} \end{cases},$$

Exercise 2

a from the definition of moment generating function (Definition 2.3.6),

$$M_Y(t) = E_Y[e^{ty}].$$

Therefore

$$\begin{aligned} M_Y(t) &= \sum_{y=0}^{\infty} e^{ty}\binom{y+r-1}{y}(1-p)^y p^r \\ &= p^r\binom{y+r-1}{y}[(1-p)e^t]^y. \end{aligned}$$

Setting $k = y$ and $a = (1-p)e^t$ in the formula of the generalized binomial series, one obtains

$$\begin{aligned} M_Y(t) &= \frac{p^r}{[1-(1-p)e^t]^r} \\ &= \left(\frac{p}{[1-(1-p)e^t]}\right)^r. \end{aligned}$$

b Since r is known, the density belongs to the exponential family,

$$f(y_1,\ldots,y_n;r,p) = f(y;r,p) = \underbrace{\prod_{i=1}^{n}\binom{y_i+r-1}{y_i}}_{h(y)}\exp\{\underbrace{\sum_{i=1}^{n} y_i}_{t(y)}\underbrace{(1-p)}_{w(p)} + \underbrace{nr\log p}_{-\kappa(p)}\},$$

and therefore $Z = \sum_{i=1}^{n} Y_i$ is sufficient and complete for p because Theorem 6.2.25.

To show that it is distributed as a negative binomial, it is sufficient to see that, due to Theorem 4.2.12,

$$\begin{aligned} M_Z(t) &= \prod_{i=1}^{n} M_{Y_i}(t) \\ &= \prod_{i=1}^{n} \left(\frac{p}{[1-(1-p)e^t]} \right)^r \\ &= \left(\frac{p}{[1-(1-p)e^t]} \right)^{nr}, \end{aligned}$$

which is the moment generating function of a Negative Binomial with parameters p and nr.

c The likelihood is

$$L(p) = \prod_{i=1}^{n} \binom{y_i + r - 1}{y_i} (1-p)^{y_i} p^r,$$

which, after removing the constants, leads to the log-likelihood

$$\ell(p) \propto \sum_{i=1}^{n} y_i \log(1-p) + nr \log p,$$

the score function

$$\ell_p(p) = -\frac{\sum_{i=1}^{n} y_i}{1-p} + \frac{nr}{p},$$

and observed information

$$\ell_{pp}(p) = -\frac{\sum_{i=1}^{n} y_i}{(1-p)^2} - \frac{nr}{p^2}.$$

Setting the score function to 0 we obtain

$$\begin{aligned} p \sum_{i=1}^{n} y_i &= nr - nrp \\ p(\sum_{i=1}^{n} y_i + nr) &= nr \end{aligned}$$

which leads to

$$\hat{p} = \frac{r}{\bar{y} + r}$$

where $\bar{y} = \sum_{i=1}^{n} y_i/n$. We know that $\hat{p}$ is a maximum because $\ell_{pp}(p)$ is always negative (all positive quantities with the minus sign).

The asymptotic variance of the MLE is equal to the inverse of the expected information. Therefore

$$\begin{aligned} i(p) &= E[-\ell_{pp}(p)] \\ &= \frac{E[\sum_{i=1}^{n} Y_i]}{(1-p)^2} + \frac{nr}{p^2} \\ &= \frac{nr(1-p)}{p(1-p)} + \frac{nr}{p^2} \\ &= \frac{nrp + nr - nrp}{p^2(1-p)} \\ &= \frac{nr}{p^2(1-p)} \end{aligned}$$

where $E[\sum_{i=1}^{n} Y_i] = nr(1-p)/p$ can be found from the results in point (b). The asymptotic variance of the MLE is therefore $\frac{p^2(1-p)}{nr}$.

d Due to the invariance property under reparamterization of the MLE,

$$\widehat{1/p} = \widehat{p^{-1}}_{MLE} = \frac{r + \bar{y}}{r}$$

and

$$\begin{aligned} E[\widehat{1/p}] &= \frac{r + E[\sum_{i=1}^{n} Y_i]/n}{r} \\ &= \frac{r + nr(1-p)/np}{r} \\ &= \frac{pr + r(1-p)}{pr} = \frac{r}{pr} = \frac{1}{p}. \end{aligned}$$

Therefore, is is unbiased.

The lowest possible variance for an unbiased estimator of $1/p$ is the Cramer-Rao lower bound, in this case, setting $\beta = 1/p$,

$$\begin{aligned} i(\beta) &= i(p(\beta)) \left(\frac{d\, p(\beta)}{d\, \beta} \right) \\ &= \frac{nr}{\frac{1}{\beta^2}\left(1 - \frac{1}{\beta}\right)} \frac{1}{\beta^4} \\ &= \frac{nr}{\frac{1}{\beta^2}\left(\frac{\beta-1}{\beta}\right)} \frac{1}{\beta^4} \\ &= \frac{nr}{\beta(\beta-1)} \end{aligned}$$

Alternatively, one could start from $\ell(\beta)$ and find $i(\beta)$.

Anyway, the asymptotic variance of $\widehat{1/p}$ is $\frac{1/p(1/p-1)}{nr}$ and it is the minimum achievable (Cramer-Rao Lower Bound) because of Theorem 10.1.12.

Exercise 3

a Since the joint density distribution belongs to the exponential family,

$$f(y_1, \ldots, y_n) = \exp\{\sum_{i=1}^{n} y_i \log \frac{p}{1-p} + n \log(1-p)\}$$

we know that $\sum_{i=1}^{n} Y_i$ is the sufficient statistic for Theorem 6.2.10.

Since it is a one-dimensional parameter exponential family and $\log(p(1-p))$ is not decreasing, the family of pmfs of $\sum_{i=1}^{n} Y_i$ have the monotone likelihood ration property

Alternatively, one could use the factorization theorem (Theorem 6.2.6) on

$$f(y_1, \ldots, y_n) = p^{\sum_{i=1}^{n} y_i}(1-p)^{n-\sum_{i=1}^{n} y_i}$$

to show that $\sum_{i=1}^{n} Y_i$ is a sufficient statistic, and show that, for $p_2 > p_1$, $\frac{f(t;p_2)}{f(t;p_2)}$ (the distribution is binomial) is monotone (increasing) in t.

b Since we know form point (a) that T is a sufficient statistics and the family of its distribution has the monotone likelihood property, we can invoke Theorem 8.3.17 (Karlin-Rubin theorem) to state that the uniform most powerful test is that that rejects the null hypothesis for $T > t_0$ at level α, where $\alpha = P(T > t_0) = \sum_{t=t_0}^{n} \binom{n}{t} p_0^t (1-p_0)^{n-t}$, as $T \sim \text{Bin}(n, p)$ with $p = p_0$ under H_0.

c from (a), setting $t(x) = \sum_{i=1}^{n} y_i$, the log-likelihood is

$$\ell(p) = t(x) \log \frac{p}{1-p} + n \log(1-p).$$

Since we know that $-2[\ell(p_0) - \ell(\hat{p})] \sim \chi_1^2$ for $n \to \infty$, the likelihood-ratio test will reject H_0 at level α when

$$-2[\ell(p_0) - \ell(\hat{p})] > \chi^2_{1;1-\alpha},$$

where $\chi^2_{1;1-\alpha}$ is the $1-\alpha$ quantile of a χ^2 distribution with 1 degree of freedom.

We can find an approximate $1-\alpha$ confidence interval by inverting this relation, resulting in

$$CI_p(1-\alpha) = \{p : -2[\ell(p) - \ell(\hat{p})] \leq \chi^2_{1;1-\alpha}\}.$$