Homework 5 solution

Trong Dat Do*

November 16, 2020

Exercise 1 **Lemma**:For $a \in (0,1)$,

$$\sum_{n=1}^{\infty} \frac{a^n}{n} = -\log(1-a),$$

and

$$\sum_{n=1}^{\infty} n a^n = \frac{a}{(1-a)^2}$$

Proof for the lemma:

$$\sum_{n=1}^{\infty} \frac{a^n}{n} = \sum_{n=1}^{\infty} \int_0^a x^{n-1} dx = \int_0^a \sum_{n=1}^{\infty} x^{n-1} dx = \int_0^a \sum_{n=1}^{\infty} \frac{1}{1-x} dx = -\log(1-a)$$

$$\sum_{n=1}^{\infty} n a^n = \sum_{n=1}^{\infty} a(a^n)' = a(\sum_{n=1}^{\infty} a^n)' = a(\frac{a}{1-a})' = \frac{a}{(1-a)^2}$$

Now, we come to the question itself.
(a)
As $x > 0$, $(1-p)^x > 0$ and $\log p < 0$, it is easy to see that

$$\mathbb{P}(X = x) = -\frac{(1-p)^x}{x \log p} > 0, x = 1, 2, 3, \ldots$$

Then, we have

$$\sum_{x=1}^{\infty} \mathbb{P}(X = x) = \sum_{x=1}^{\infty} -\frac{(1-p)^x}{x \log p} = \frac{1}{\log p} \times \log(1-(1-p)) = 1(\text{ We use the lemma here.})$$

Thus, this is a proper probability distribution.

(b)

*Ph.D. Student, Department of Statistics, University of Michigan

It is easy to find that

$$-\frac{(1-p)^x}{x\log p}$$

is a decreasing function w.r.t. x for $x > 0$, because $(1-p)^x > 0$ and is decreasing with w.r.t. x, $\frac{1}{x} > 0$ and is decreasing with w.r.t. x and $-\frac{1}{\log p} > 0$. Thus, the most likely number of sights of the jungle animal is 1.

(c)

$$\mathbb{E}[X] = \sum_{x=1}^{\infty} -x\frac{(1-p)^x}{x\log p} = \sum_{x=1}^{\infty} -\frac{(1-p)^x}{\log p} = -\frac{(1-p)}{p\log p}$$

$$\mathbb{E}[X^2] = \sum_{x=1}^{\infty} -x\frac{(1-p)^x}{\log p} = -\frac{1}{\log p}\frac{1-p}{p^2} = -\frac{1-p}{p^2\log p}(\text{ We use the lemma here})$$

Thus the variance is

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = -\frac{1-p}{p^2\log p} - (\frac{(1-p)}{p\log p})^2 = -\frac{1-p}{p^2\log p}(1+\frac{1-p}{\log p})$$

Exercise 2 (a)

$$
\begin{aligned}
M_{\underline{X}}(\mathbf{t}) &= \mathbb{E}[\exp(\mathbf{t}^t \underline{X})] = \mathbb{E}[\exp(\sum_{i=1}^{p} t_i X_i)] \\
&= \mathbb{E}[\prod_{i=1}^{p} \exp(t_i X_i)] = \prod_{i=1}^{p} \mathbb{E}[\exp(t_i X_i)] = \prod_{i=1}^{p} M_{X_i}(t_i)
\end{aligned}
$$

(b)

For a standard normal r.v. X, the m.g.f of it is

$$M_X(t) = \exp(\frac{t^2}{2})$$

Then, according to the result in (a), we have

$$M_{\underline{X}}(\mathbf{t}) = \exp(\sum_{i=1}^{p} t_i^2/2)$$

(c)

We only need to calculate $M_{\mu,\Sigma}(.)$ here.

$$
\begin{aligned}
M_{\underline{Y}}(\mathbf{t}) &= \mathbb{E}[\exp(\mathbf{t}^t \underline{Y})] = \mathbb{E}[\exp(\mathbf{t}^t(\mu + B\underline{X}))] \\
&= \exp(\mathbf{t}^t \mu)\mathbb{E}[\exp(\mathbf{t}^t B\underline{X})] = \exp(\mathbf{t}^t \mu) M_{\underline{X}}((B^t)\mathbf{t}) \\
&= \exp(\mathbf{t}^t \mu) \exp(\mathbf{t}^t (B)(B^t)\mathbf{t}/2) \\
&= \exp(\mathbf{t}^t \mu + \mathbf{t}^t \Sigma \mathbf{t}/2)
\end{aligned}
$$

which only rely on μ and Σ.

(d)

Define

$$D_{Y,1}(\mathbf{t}) = \frac{\partial M_{\underline{Y}}(\mathbf{t})}{\partial \mathbf{t}} = (\mu + \Sigma\mathbf{t}) \exp(\mathbf{t}^t \mu + \mathbf{t}^t \Sigma \mathbf{t}/2)$$

$$D_{Y,2}(\mathbf{t}) = \frac{\partial M_{\underline{Y}}(\mathbf{t})^2}{\partial^2 \mathbf{t}} = (\mu + \Sigma\mathbf{t})(\mu + \Sigma\mathbf{t})^T \exp(\mathbf{t}^t \mu + \mathbf{t}^t \Sigma \mathbf{t}/2) + \Sigma \exp(\mathbf{t}^t \mu + \mathbf{t}^t \Sigma \mathbf{t}/2)$$

It is easy to check that

$$D_{Y,1}(\mathbf{0}) = \mu$$

$$D_{Y,2}(\mathbf{0}) - D_{Y,1}(\mathbf{0}) D_{Y,1}(\mathbf{0})^T = \Sigma$$

Now, assume there are two r.v.'s Y_1 and Y_2 equipped with μ_1, Σ_1 and μ_2, Σ_2. If $M_{\mu_1,\Sigma_1} = M_{\mu_2,\Sigma_2}$, then we must have $D_{Y_1,1}(\mathbf{0}) = D_{Y_2,1}(\mathbf{0})$ and $D_{Y_1,2}(\mathbf{0}) - D_{Y_1,1}(\mathbf{0}) D_{Y_1,1}(\mathbf{0})^T = D_{Y_2,2}(\mathbf{0}) - D_{Y_2,1}(\mathbf{0}) D_{Y_2,1}(\mathbf{0})^T$.

However,

$$
\begin{aligned}
&(D_{Y_1,1}(\mathbf{0}), D_{Y_1,2}(\mathbf{0}) - D_{Y_1,1}(\mathbf{0}) D_{Y_1,1}(\mathbf{0})^T) = (\mu_1, \Sigma_1) \\
&\neq (\mu_2, \Sigma_2) = (D_{Y_2,1}(\mathbf{0}), D_{Y_2,2}(\mathbf{0}) - D_{Y_2,1}(\mathbf{0}) D_{Y_2,1}(\mathbf{0})^T)
\end{aligned}
$$

which is a contradiction. Thus, $M_{\mu_1,\Sigma_1} \neq M_{\mu_2,\Sigma_2}$ and Y_1, Y_2 have different distributions.

(e)

For the $\underline{X}$ in (a), it is easy to see that the density for that is

$$f_{\underline{X}}(\underline{x}) = \prod_{i=1}^{p} f_{X_i}(x_i) = \prod_{i=1}^{p} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x_i^2}{2}) = \frac{1}{(\sqrt{2\pi})^p} \exp(-\underline{x}^T \underline{x}/2)$$

Now, as $\underline{Y} = \mu + B\underline{X}$, we have

$$|\frac{\partial \underline{X}}{\partial \underline{Y}}| = |B^{-1}| = |\Sigma|^{-1/2}$$

Thus,

$$\begin{aligned} f_{\underline{Y}}(\underline{y}) &= f_{\underline{X}}(B^{-1}(\underline{y} - \mu))|\Sigma|^{-1/2} \\ &= \frac{1}{(\sqrt{2\pi})^p} \exp(-(B^{-1}(\underline{y} - \mu))^T (B^{-1}(\underline{y} - \mu))/2)\Sigma|^{-1/2} \\ &= \frac{1}{(2\pi)^{p/2}\Sigma|^{1/2}} \exp(-(\underline{y} - \mu)^T \Sigma^{-1} (\underline{y} - \mu)/2) \end{aligned}$$

Exercise (4.33) We have

$$
\begin{aligned}
MGF_H(t) &= Ee^{Ht} \\
&= E[E[e^{Ht}|N]] \qquad \text{(iterated expectation)} \\
&= E[E[e^{(\sum_{i=1}^{N} X_i)t}|N]] \qquad \text{(from definition)} \\
&= E[E[\prod_{i=1}^{N} e^{X_i t}|N]] \\
&= E[\prod_{i=1}^{N} E[e^{X_i t}|N]] \qquad \text{(conditionally independent)} \\
&= E\left[\left(\frac{\log(1-(1-p)e^t)}{\log p}\right)^N\right] \\
&= \sum_{n=0}^{\infty}\left(\frac{\log(1-(1-p)e^t)}{\log p}\right)^n \lambda^n e^{-\lambda}/n! \\
&= e^{-\lambda}\sum_{n=0}^{\infty}\left(\frac{\log(1-(1-p)e^t)}{\log p}\lambda\right)^n /n! \\
&= e^{-\lambda}e^{\lambda\log(1-(1-p)e^t)/(\log p)} \\
&= e^{r\log p}e^{-r\lambda\log(1-(1-p)e^t)} \qquad \text{(set } r=-\lambda/\log p\text{)} \\
&= \left(\frac{p}{(1-(1-p)e^t)}\right)^r,
\end{aligned}
$$

which is the MGF of Negative Binomial distribution with parameter r and $1-p$. □

Exercise (4.34) (a) We just integral out P to get the distribution of X.

$$\begin{aligned}
P(X=x) &= E[P(X=x|P)] \\
&= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} P(X=x|P=p) \\
&= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} \binom{n}{x} p^x (1-p)^{n-x} \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \binom{n}{x} \int_0^1 p^{x+\alpha-1}(1-p)^{n-x+\beta-1} \\
&= \binom{n}{x} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+x)\Gamma(\beta+n-x)}{\Gamma(\alpha+\beta+n)}.
\end{aligned}$$

(b) We do the same thing to get the pmf of X

$$\begin{aligned}
P(X=x) &= E[P(X=x|P)] \\
&= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} P(X=x|P=p) \\
&= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} \binom{n+x-1}{x} p^x (1-p)^r \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \binom{n+x-1}{x} \int_0^1 p^{x+\alpha-1}(1-p)^{r+\beta-1} \\
&= \binom{n+x-1}{x} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+x)\Gamma(\beta+r)}{\Gamma(\alpha+\beta+n+r)}.
\end{aligned}$$

For the expectation, we have

$$\begin{aligned}
EX &= E[E[X|P]] \\
&= E\left[\frac{Pr}{1-P}\right] \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 r p^{\alpha}(1-p)^{\beta-2} dp \\
&= \frac{\Gamma(\alpha+1)\Gamma(\beta-1)}{\Gamma(\alpha)\Gamma(\beta)} r \\
&= \frac{\alpha}{\beta-1} r
\end{aligned}$$

Similarly, for the variance

$$\begin{aligned}
\operatorname{Var} X &= \operatorname{Var}(E[X|P]) + E[\operatorname{Var}(X|P)] \\
&= \operatorname{Var}\left(\frac{Pr}{1-P}\right) + E\left[\frac{Pr}{(1-P)^2}\right] \\
&= r^2 \left(\frac{\Gamma(\alpha+2)\Gamma(\beta-2)}{\Gamma(\alpha)\Gamma(\beta)} - \left(\frac{\Gamma(\alpha+1)\Gamma(\beta-1)}{\Gamma(\alpha)\Gamma(\beta)}\right)^2\right) + r\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+1)\Gamma(\beta-2)}{\Gamma(\alpha+\beta-1)} \\
&= r^2 \left(\frac{\alpha(\alpha+1)}{(\beta-1)(\beta-2)}\right)
\end{aligned}$$

□

Exercise (4.36) (a)

$$EY = \sum_{i=1}^{n} EX_i = \sum_{i=1}^{n} E[E[X_i|P_i]] = \sum_{i=1}^{n} EP_i = \frac{\alpha}{\alpha+\beta}.$$

(b) Because of the independence,

$$\begin{aligned}
\operatorname{Var} Y &= \sum_{i=1}^{n} \operatorname{Var} X_i \\
&= \sum_{i=1}^{n} \operatorname{Var} E[X_i|P_i] + E[\operatorname{Var}(X_i|P_i)] \\
&= \sum_{i=1}^{n} \operatorname{Var} P_i + E[P_i(1-P_i)] \\
&= \sum_{i=1}^{n} EP_i^2 - (EP_i)^2 + EP_i - EP_i^2 \\
&= \sum_{i=1}^{n} EP_i - (EP_i)^2 \\
&= n\left(\frac{\alpha}{\alpha+\beta} - \frac{\alpha^2}{(\alpha+\beta)^2}\right) \\
&= n\frac{\alpha\beta}{(\alpha+\beta)^2}.
\end{aligned}$$

To find the distribution of Y, notice that Y is the sum of i.i.d. X_i, and $P(X_i = 1) = E[P(X_i = 1|P_i)] = \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} p dp = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+\beta+1)}{\Gamma(\alpha+1)\Gamma(\beta)} = \frac{\alpha}{\alpha+\beta}$. Hence Y has binomial distribution.

(c) Again because of the independence of X_i's,

$$EY = \sum_{i=1}^{k} EX_i, \operatorname{Var} Y = \sum_{i=1}^{k} \operatorname{Var} X_i.$$

Now, similar to above,

$$EX_i = E[X_i|P_i] = n_i \frac{\alpha}{\alpha+\beta},$$

$$\begin{aligned}
\operatorname{Var} X_i = \operatorname{Var} E[X_i|P_i] + E \operatorname{Var}[X_i|P_i] &= \operatorname{Var}(n_i P_i) + E(n_i P_i(1-P_i)) \\
&= n_i \frac{\alpha\beta(\alpha+\beta+n_i)}{(\alpha+\beta)^2(\alpha+\beta+1)},
\end{aligned}$$

where we use the fact that $EP_i^2 = \dfrac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$. □

Exercise (4.40) (1) We have

$$1/C = \int_{x,y\geq 0, x+y\leq 1} x^{a-1}y^{b-1}(1-x-y)^{c-1}dxdy.$$

A way to calculate this integral is to use the Gamma integral. Using change of variables $x = \dfrac{z_1}{z_1+z_2+z_3}, y = \dfrac{z_2}{z_1+z_2+z_3}, t = z_1+z_2+z_3$, we have

$$\begin{aligned}
\Gamma(a)\Gamma(b)\Gamma(c) &= \int_0^\infty dz_1 z_1^{a-1}e^{-z_1}\int_0^\infty dz_2 z_2^{a-1}e^{-z_2}\int_0^\infty dz_3 z_3^{c-1}e^{-z_3}\\
&= \int_{0\leq x,y,1\geq x+y} x^{a-1}y^{b-1}(1-x-y)^{c-1}dxdy\int_0^\infty t^{a+b+c-1}e^{-t}dt\\
&= \int_{0\leq x,y,1\geq x+y} x^{a-1}y^{b-1}(1-x-y)^{c-1}dxdy\Gamma(a+b+c).
\end{aligned}$$

where we use the Jacobian $dz_1dz_2dz_3 = t^2dxdydt$.

(2) We prove that $X \sim Beta(a, b+c)$. Similarly we would have $Y \sim beta(b, a+c)$.

$$\begin{aligned}
f(x) &= \int_{0\leq y\leq 1-x} f(x,y)dy\\
&\propto x^{a-1}\int_0^{1-x} y^{b-1}(1-y-x)^{c-1}dy\\
&= x^{a-1}(1-x)^{b+c-2}\int_0^{1-x}\left(\frac{y}{1-x}\right)^{b-1}\left(\frac{1-x-y}{1-x}\right)^{c-1}dy\\
&= x^{a-1}(1-x)^{b+c-2}\int_0^1 (z)^{b-1}(1-z)^{c-1}dz(1-x) \quad (\text{change } z = y/(1-x))\\
&\propto x^{a-1}(1-x)^{b+c-1}.
\end{aligned}$$

where f $\propto$ g means f/g equals to a constant which does not depend on x, y.

(3) Using the results above, we have

$$\begin{aligned}
f_{Y|X=x}(y) &= \frac{f(x,y)}{f(x)}\\
&\propto x^{a-1}y^{b-1}(1-x-y)^{c-1}/(x^{a-1}(1-x)^{b+c-1})\\
&\propto \left(\frac{y}{1-x}\right)^{b-1}\left(\frac{1-x-y}{1-x}\right)^{c-1}\frac{1}{1-x}.
\end{aligned}$$

Hence using the change of variable formula we have

$$f_{Y/(1-X)|X}(t|x) \propto t^{b-1}(1-t)^{c-1},$$

which does not depend on x. Hence $Y/(1-X)$ and X are independent, and also $Y/(1-X) \sim Beta(b, c)$.

(4) We have

$$\begin{aligned}
EXY &= \frac{\gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} \int_{0\le x,y,x+y\le 1} x^a y^b (1-x-y)^{c-1} dxdy \\
&= \frac{\gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} \times \frac{\Gamma(a+1)\Gamma(b+1)\Gamma(c)}{\gamma(a+b+c+2)} \\
&= \frac{ab}{(a+b+c)(a+b+c+1)}.
\end{aligned}$$

Similarly we have

$$\begin{aligned}
E(XY)^2 &= \frac{\gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} \int_{0\le x,y,x+y\le 1} x^{a+1} y^{b+1} (1-x-y)^{c-1} dxdy \\
&= \frac{\gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} \times \frac{\Gamma(a+2)\Gamma(b+2)\Gamma(c)}{\gamma(a+b+c+4)} \\
&= \frac{a(a+1)b(b+1)}{(a+b+c)(a+b+c+1)(a+b+c+2)(a+b+c+3)}.
\end{aligned}$$

and $\mathrm{Var}(XY) = E(XY)^2 - (E(XY))^2$ $\square$

Exercise (4.47) (1) We have

$$\begin{aligned} f_Z(z) &= f_Z(z|XY > 0)P(XY > 0) + f_Z(z|XY < 0)P(XY < 0) \\ &= f_X(z)P(XY > 0) + f_X(-z)P(XY < 0) \\ &= f_X(z), \end{aligned}$$

because X is symmetric.

(2) If Y and Z are multivariate normal, then because its support is all $\mathbb{R}^2$, we must have $P(ZY < 0) > 0$. But from the definition of Z we have $ZY = XY > 0$ if $XY > 0$ and $ZY = -XY > 0$ if $XY < 0$. Hence $P(YZ > 0) = 1$, which implies $P(XZ < 0) = 0$, a contradiction. $\square$

Exercise 3 (a)

$$F(x) = \int_0^x f(u)du \geq \int_0^x f(x)du = f(x)\int_0^x du = xf(x)$$

$$\mu = \int_0^M uf(u)du \leq \int_0^M F(u)du$$

(b)

To prove $m \leq \frac{M}{2}$, we only need to prove that $\mathbb{P}(X \leq \frac{M}{2}) \geq 0.5$ and this is equivalent to prove $\mathbb{P}(X \leq \frac{M}{2}) \geq \mathbb{P}(X \geq \frac{M}{2})$.

Then we have

$$\mathbb{P}(X \leq \frac{M}{2}) = \int_0^{\frac{M}{2}} f(u)du \geq \int_0^{\frac{M}{2}} f(u + \frac{M}{2})du = \int_{\frac{M}{2}}^M f(u)du = \mathbb{P}(X \geq \frac{M}{2})$$

Thus, $m \leq \frac{M}{2}$.

According to the result in (a), we have

$$\int_0^M (1 - F(x))dx = \mu \leq \int_0^M F(x)dx$$

Thus,

$$M = \int_0^M 1dx \leq 2\int_0^M F(x)dx$$

Therefore,

$$\mu = \int_0^M (1 - F(x))dx = M - \int_0^M F(x)dx \leq \frac{M}{2}$$

(c)

According to the steps in (b), we know that if we have $m = \frac{M}{2}$, we must have

$$\int_0^{\frac{M}{2}} f(u)du = \int_0^{\frac{M}{2}} f(u + \frac{M}{2})du$$

Thus,

$$\int_0^{\frac{M}{2}} f(u) - f(u + \frac{M}{2})du = 0$$

Denote $h(u) = f(u) - f(u + \frac{M}{2}), u \in [0, \frac{M}{2}]$. As $f(u)$ is non-increasing and continuous, we have $h(u) \geq 0$ and continuous. But as $\int_0^{\frac{M}{2}} h(u)du = 0$, we have $h(u) = 0, u \in [0, \frac{M}{2}]$, that is $f(u) = f(u + \frac{M}{2}), u \in [0, \frac{M}{2}]$.

For $\forall y \in [0, \frac{M}{2}]$,

$$f(y + \frac{M}{2}) = f(y) \geq f(\frac{M}{2}) \geq f(y + \frac{M}{2}).$$

Therefore $f(y) = f(y + \frac{M}{2}) = f(\frac{M}{2})$, which means $f(u)$ is a constant.

If $\mu = \frac{M}{2}$, according to the steps in (b), we must have

$$\int_0^M F(x) = \frac{M}{2}$$

This means we have $\mu \leq \int_0^M F(x)dx$. Then according to the steps in (a), we must have

$$\int_0^x f(u)du \geq \int_0^x f(x)du$$

We can now set $x = M$ and it will be

$$\int_0^M f(u) - f(1)du = 0$$

As $f(u) - f(M)$ is non-negative and continuous, we have $f(u) = f(M)$, which means $f(u)$ is a constant. $\square$