統方 HW3 數據 114 徐仁瓏 RE6121011

a) ⓒ 令 $X_1, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

$E(X) = \mu \overset{\text{set}}{=} \bar{X}$

$Var(X) = \sigma^2 \overset{\text{set}}{=} M_2$ $\qquad M_r = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^r, \ r = 2, \cdots$

$$\Rightarrow \begin{cases} \hat{\mu}_{MME} = \bar{X} \\ \hat{\sigma}^2_{MME} = M_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \end{cases}$$

ⓓ 令 $X_1, \cdots, X_n \overset{iid}{\sim} Exp(\lambda)$

$E(X) = \frac{1}{\lambda} \overset{\text{set}}{=} \bar{X}$

$$\Rightarrow \hat{\lambda}_{MME} = \frac{1}{\bar{X}} = \frac{n}{\sum_{i=1}^{n} X_i}$$

	Dataset 1	Dataset 2
Normal	$\hat{\mu} \doteq 0.0247$	$\hat{\mu} \doteq 19.2467$
	$\hat{\sigma}^2 \doteq 0.7067$	$\hat{\sigma}^2 \doteq 25.9126$
Exponential	$\hat{\lambda} \doteq 40.4680$	$\hat{\lambda} \doteq 0.051957$
Laplace	$\hat{\mu} \doteq 0.0247$	$\hat{\mu} \doteq 19.2467$
	$\hat{b} \doteq 0.5944$	$\hat{b} \doteq 3.5995$
Gamma	$\hat{\alpha} \doteq 0.000864$	$\hat{\alpha} \doteq 14.295566$
	$\hat{\lambda} \doteq 0.034965$	$\hat{\lambda} \doteq 0.742754$

ⓔ 令 $X_1, \cdots, X_n \overset{iid}{\sim} DE(\mu, b)$

$E(X) = \mu \overset{\text{set}}{=} \bar{X}$

$Var(X) = 2b^2 \overset{\text{set}}{=} M_2$

$$\Rightarrow \begin{cases} \hat{\mu}_{MME} = \bar{X} \\ \hat{b}_{MME} = \sqrt{\frac{M_2}{2}} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{2n}} \end{cases}$$

ⓕ 令 $X_1, \cdots, X_n \overset{iid}{\sim} Gamma(\alpha, \lambda)$

$E(X) = \frac{\alpha}{\lambda} \overset{\text{set}}{=} \bar{X} \quad \cdots ①$

$Var(X) = \frac{\alpha}{\lambda^2} \overset{\text{set}}{=} M_2 \quad \cdots ②$

$\Rightarrow \frac{①}{②}: \lambda = \frac{\bar{X}}{M_2}$

$\alpha = \lambda \cdot \bar{X} = \frac{\bar{X}^2}{M_2}$

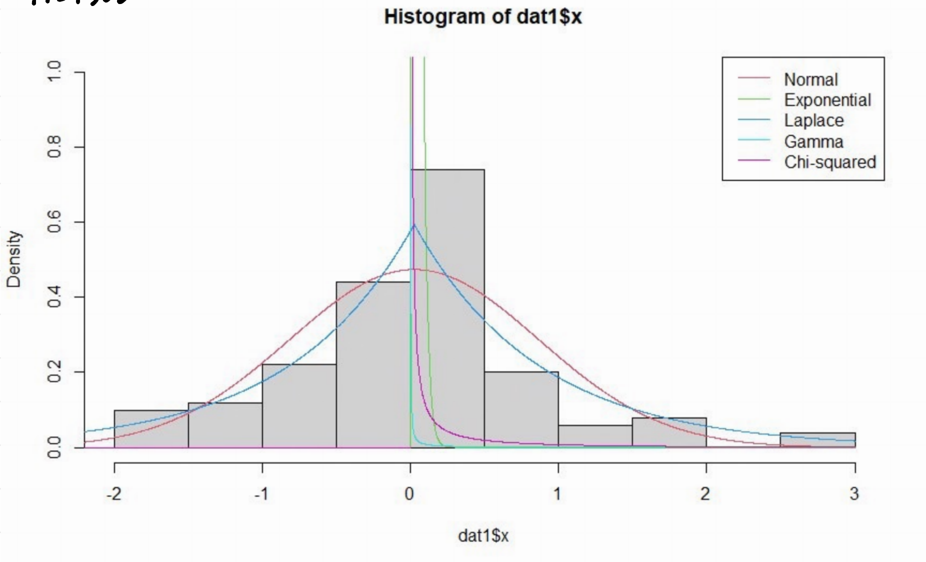
$$\Rightarrow \begin{cases} \hat{\alpha}_{MME} = \frac{\bar{X}^2}{M_2} = \frac{n \cdot \bar{X}^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \\ \hat{\lambda}_{MME} = \frac{\bar{X}}{M_2} = \frac{n \cdot \bar{X}}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \end{cases}$$

ⓖ 令 $X_1, \cdots, X_n \overset{iid}{\sim} \chi^2(k)$

$E(X) = k \overset{\text{set}}{=} \bar{X}$

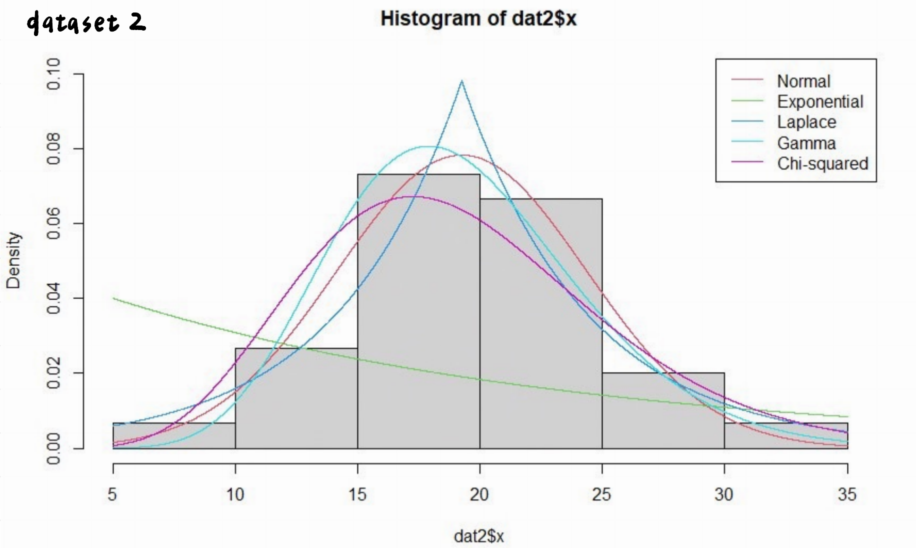
$$\Rightarrow \hat{k}_{MME} = \bar{X}$$

b) data set 1



從上圖可知，只有 Normal 和 Laplace 分配 fit 得比較好，
因為 Exponential，Gamma，Chi-squared 的範圍都必須 >0，
所以與此 data 較不符合。

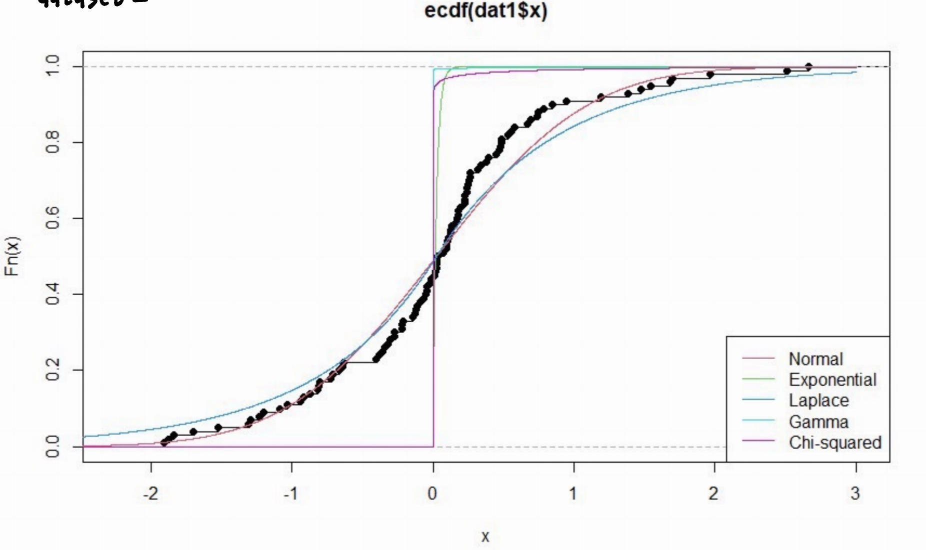
dataset 2



從上圖可知，除了 Exponential 之外，其它分配都 fit 得不錯，
因為 Exponential 的眾數永遠在 0，因此與此 data 不符合。

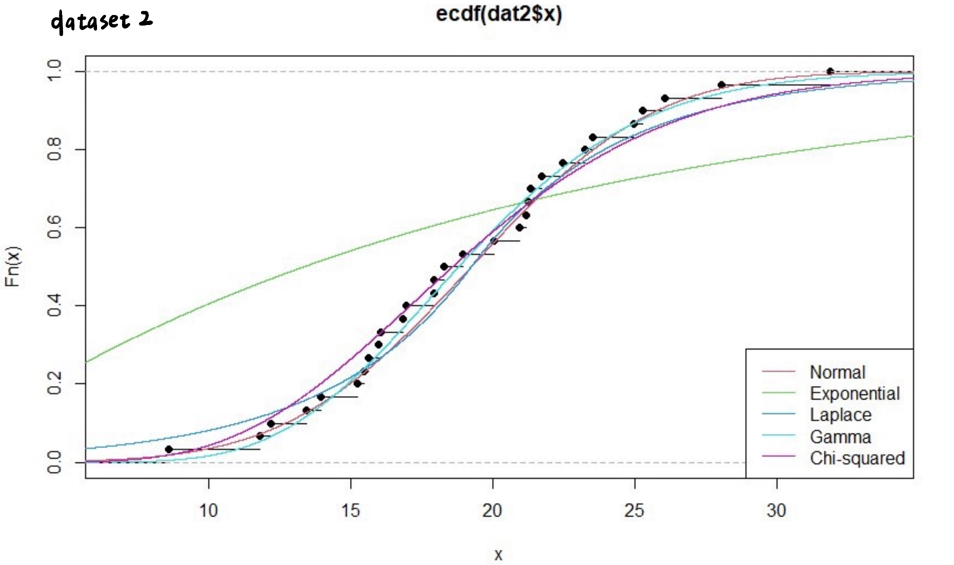
c)

dataset 1



從上圖可知，只有 Normal 和 Laplace 分配 fit 得比較好，但在 [0,1] 這段 fit 比較不好，可能是因為此 data 並沒有那麼對稱。

dataset 2



從上圖可知，除了 Exponential 之外，其它分配都 fit 得不錯，其中，在尾端的部分，似乎是 Normal 和 Gamma 較佳。

d) dataset 1

	Normal	Exponential	Laplace	Gamma	Chisq
KS test	0.2791435	0e+00	0.3350083	0e+00	0e+00
AD test	0.1605055	6e-06	0.1226789	6e-06	6e-06
CVMtest	0.1386345	0e+00	0.1429036	0e+00	0e+00

從上表可知，在三种檢定下，Exponential，Gamma，Chi-squared 的 p-value 皆小於 0.05，代表拒絕 H_0，表示我們有足夠的證據證明此資料不服從這三种分配。而 Normal 和 Laplace 的 p-value 皆大於 0.05，代表不拒絕 H_0，表示我們可以假設此資料服從這兩個分配。

dataset 2

	Normal	Exponential	Laplace	Gamma	Chisq
KS test	0.9824112	3.000051e-05	0.9360950	0.9539047	0.9360950
AD test	0.9983586	2.091198e-04	0.8617125	0.9991130	0.8785581
CVMtest	0.9897424	8.917683e-05	0.8814181	0.9955509	0.8564987

從上表可知，在三种檢定下，只有 Exponential 的 p-value 皆小於 0.05，代表拒絕 H_0，表示我們有足夠的證據證明此資料不服從此分配。而 Normal，Laplace，Gamma 和 Chi-squared 的 p-value 皆遠大於 0.05，代表不拒絕 H_0，表示我們可以假設此資料服從這四個分配。其中，Normal 和 Gamma 的 p-value 較大，又看过上面兩張 PDF 和 CDF 的图後，我認為此兩個分配與此筆資料較為相符。