Solutions_HW2

Wei-Rou,Xu

2023-11-15

1. Give the sample space (domain), probability density function, mean, and variance of the following distributions:

(a) Poisson distribution with rate $\lambda = 3$:

$$f(x) = \frac{e^{-3}3^x}{x!}, x = 0, 1, 2, ...$$
$$E(X) = 3,$$
$$Var(X) = 3.$$

(b) Geometric distribution with probability $p = 1/3$:

$$f(x) = (\frac{1}{3})(\frac{2}{3})^{x-1}, x = 1, 2, 3, ...$$
$$E(X) = 3,$$
$$Var(X) = 6.$$
$$f(x) = (\frac{1}{3})(\frac{2}{3})^{x}, x = 0, 1, 2, ...$$
$$E(X) = 2,$$
$$Var(X) = 6.$$

(c) $X \sim N(0, 1)$:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{x^2}{2}\right\}, \ -\infty < x < \infty.$$
$$E(X) = 0,$$
$$Var(X) = 1.$$

(d) Exponential distribution with rate $\lambda = 2$:

$$f(x) = 2e^{-2x}, \ x > 0.$$
$$E(X) = \frac{1}{2},$$
$$Var(X) = \frac{1}{4}.$$

(e) Laplace distribution with location parameter 0 and scale parameter 1/2:

$$f(x) = e^{-2|x|}, \ -\infty < x < \infty.$$
$$E(x) = 0,$$
$$Var(X) = \frac{1}{2}.$$

(f) Gamma distribution with shape parameter 10 and rate parameter 1/2:

$$f(x) = \frac{(1/2)^{10}}{\Gamma(10)} x^9 e^{-x/2}, x > 0.$$

$$E(X) = 20,$$
$$Var(X) = 40.$$

(g) Chi-square distribution with degree of freedom 20:

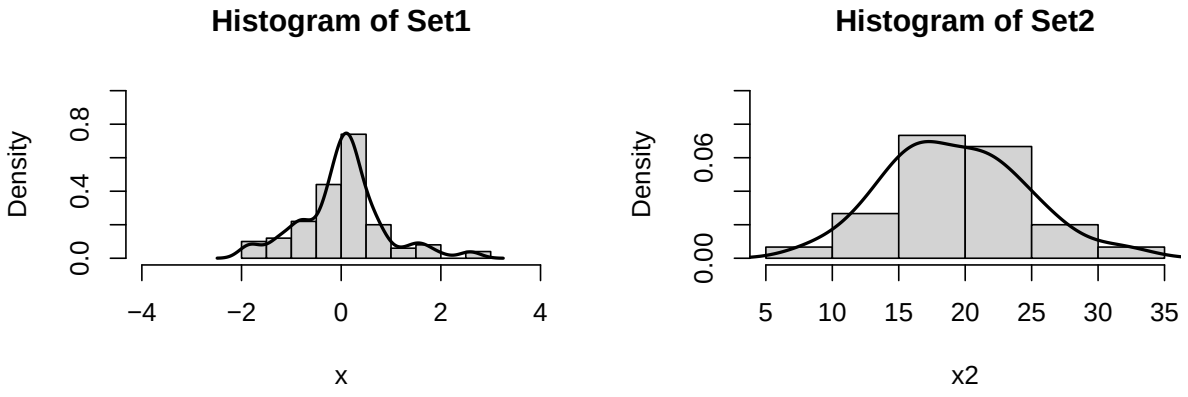
$$f(x) = \frac{1}{\Gamma(10)2^{10}} x^9 e^{-x/2}, x > 0.$$

$$E(X) = 20,$$
$$Var(X) = 40.$$

2. Given two datasets,

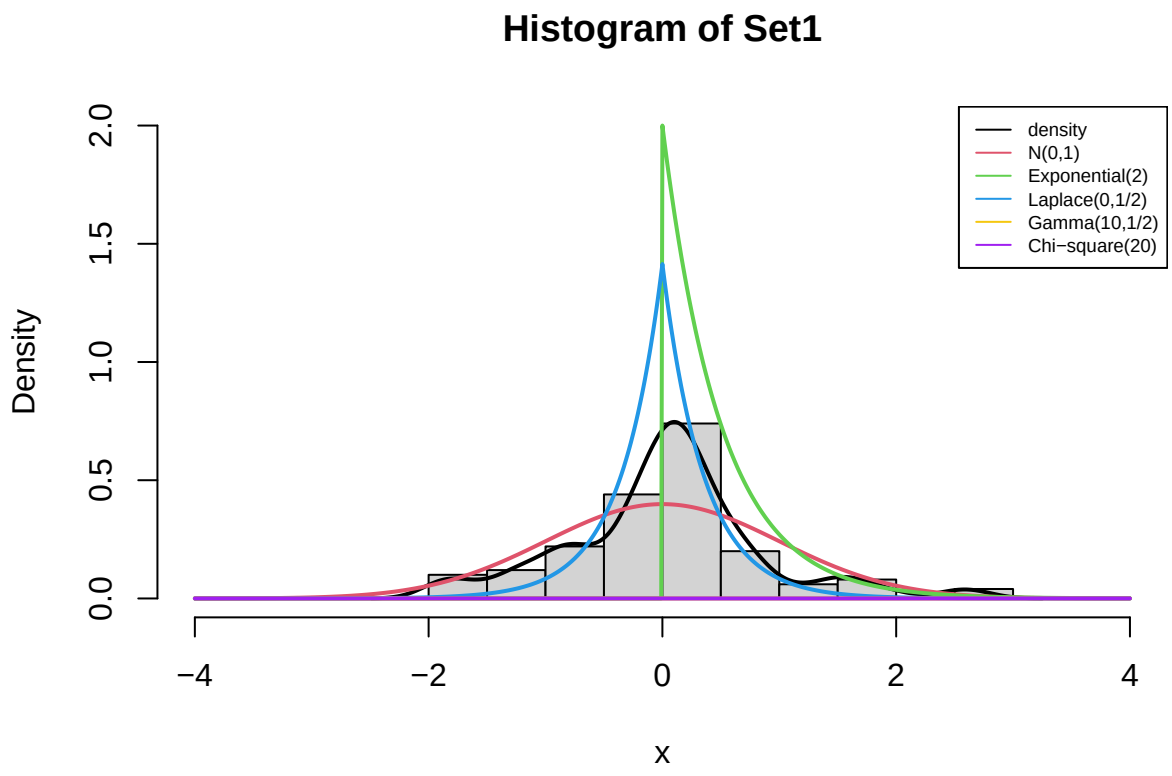
(a) Please provide the histograms of two datasets.

```
set1=read.csv("C:/Set1.csv")
x=set1$x
par(mfrow = c(1, 2))
hist(x, probability = TRUE, xlim = c(-4, 4), ylim = c(0, 1),main="Histogram of Set1")
lines(density(x), col = 1, lwd = 2)

set2=read.csv("C:/Set2.csv")
x2=set2$x
hist(x2, probability = TRUE, xlim = c(5, 35), ylim = c(0, 0.10),main="Histogram of Set2")
lines(density(x2), col = 1, lwd = 2)
```



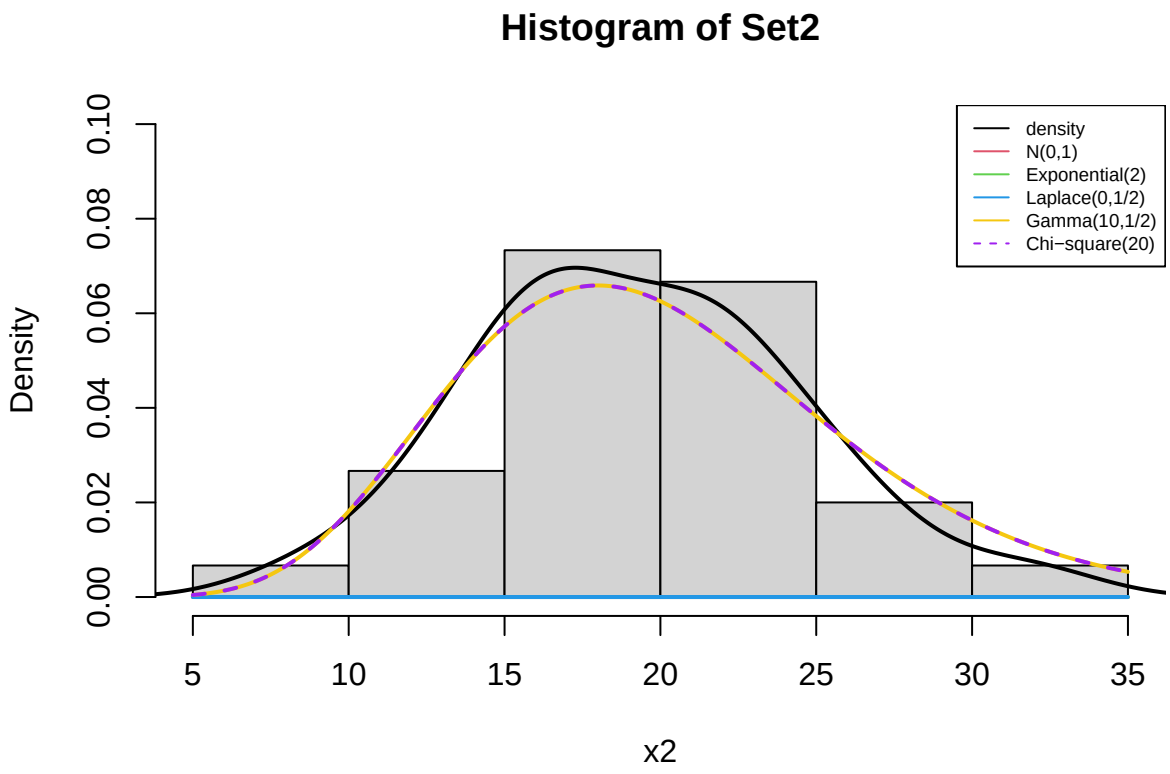
(b) For each dataset, add the probability density functions of the given distributions (c)-(g) in Question 1 to the figures in Question 2(a). Tying to select more suitable distributions to the data based on your opinion.

```r
xx <- seq(-4,4, 0.01)
hist(x, probability = TRUE, xlim = c(-4, 4), ylim = c(0, 2),main="Histogram of Set1")
lines(density(x), col = 1, lwd = 2)
lines(xx, dnorm(xx, 0, 1), col = 2, lwd = 2)
lines(xx, dexp(xx, 2), col = 3, lwd = 2)
library(jmuOutlier)
lines(xx, dlaplace(xx,0,1/2), col = 4, lwd = 2)
lines(xx, dgamma(xx,shape=10,rate =0.5), col = 7, lwd = 2)
lines(xx, dchisq(xx,20), col = "Purple", lwd = 2)

legend("topright", c("density", "N(0,1)", "Exponential(2)",
                     "Laplace(0,1/2)","Gamma(10,1/2)","Chi-square(20)"),
       lty = c(1, 1), col = c(1,2,3,4,7,"Purple"), cex = 0.6)
```




```r
hist(x2, probability = TRUE, xlim = c(5, 35), ylim = c(0, 0.10),main="Histogram of Set2")
lines(density(x2), col = 1, lwd = 2)
xx2 <- seq(5,35, 0.01)
lines(xx2, dnorm(xx2, 0, 1), col = 2, lwd = 2)
lines(xx2, dexp(xx2, 2), col = 3, lwd = 2)
library(jmuOutlier)
lines(xx2, dlaplace(xx2,0,1/2), col = 4, lwd = 2)
lines(xx2, dgamma(xx2,shape=10,rate =0.5), col = 7, lwd = 2)
lines(xx2, dchisq(xx2,20), col = "Purple", lwd = 2,lty =2)
legend("topright", c("density", "N(0,1)", "Exponential(2)",
                     "Laplace(0,1/2)","Gamma(10,1/2)","Chi-square(20)"),
```

```
       lty = c(1, 1,1,1,1,2), col = c(1,2,3,4,7,"Purple"), cex = 0.6)
```



From the figures, I suggest that the laplace distribution is for set 1 and the gamma distribution (or chi-squared distribution) for set 2.