Fitting to Data (Exercise for Homework)

TLS

November 13, 2022

E3

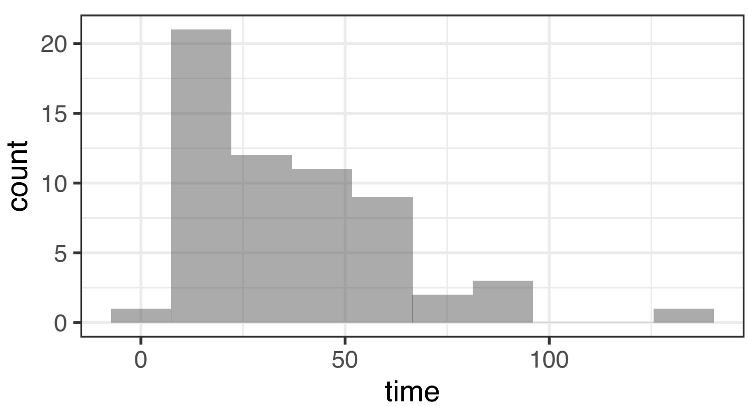
Part (a)

I look at the .csv file containing times between made baskets in a men's Calvin-Kalamazoo basketball game.

```
bballGame <- read.csv("http://scofield.site/teaching/data/csv/stob/scores.csv")
head(bballGame)
```

```
##   time
## 1   37
## 2   59
## 3  140
## 4   80
## 5   18
## 6   24
```

```
gf_histogram(~time, data=bballGame, bins=10)
```




```
nrow(bballGame)
```

```
## [1] 60
```

Part (b)

In fitting a distribution using the method of moments, we need $\overline{x}$ and $v = \dfrac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x})^2$. We calculate these as

```
xbar = mean(~time, data=bballGame); xbar
```

```
## [1] 36.85
```

```
v = var(~time, data=bballGame) * 59/60; v
```

```
## [1] 647.3275
```

I will fit this data with a gamma distribution `Gamma`(α, λ). When we studied the method of moments, we inverted the relationships for estimates $\hat{\alpha}$, $\hat{\lambda}$ of the parameters:

$$\overline{x} = \frac{\hat{\alpha}}{\hat{\lambda}}, \quad v = \frac{\alpha}{\lambda^2} \qquad \Rightarrow \qquad \hat{\lambda} = \frac{\overline{x}}{v}, \quad \hat{\alpha} = \frac{\overline{x}^2}{v}.$$

We compute these numbers:

```
hatLambda = xbar / v; hatLambda
```

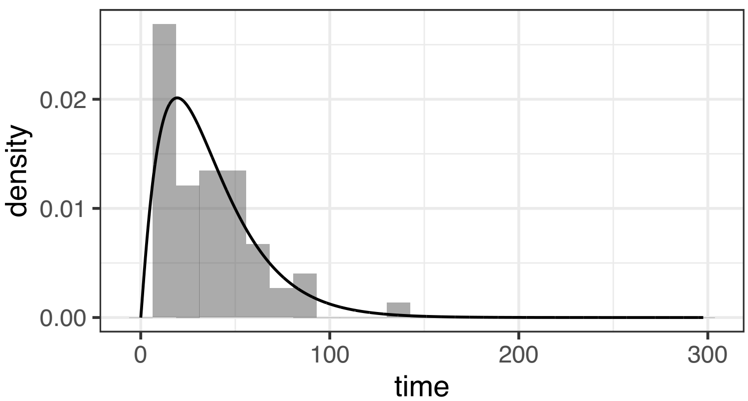
```
## [1] 0.05692636
```

```
hatAlpha = xbar^2 / v; hatAlpha
```

```
## [1] 2.097736
```

Plotting the histogram and the Gamma-distribution with these estimated parameters, we have

```
gf_dhistogram(~time, data=bballGame) |>
  gf_dist("gamma", params=c(hatAlpha, hatLambda))
```



Part (c)

Here is code that is slightly modified from the textbook for base kernels and kernel fitting:

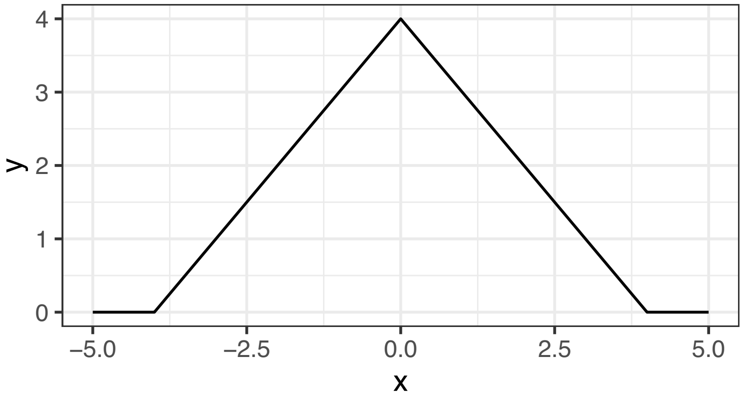
```
K1 <- function(x) {return(as.numeric(-1 < x & x < 1))}
K2 <- function(x) {return( (4-abs(x)) * (abs(x) < 4) )}
K3 <- function(x) {return( (1-x^2) * (abs(x) < 1) )}
K4 <- dnorm
K5 <- function(x) {return(dnorm(x, 0, 2))}

kde <- function(data, kernel=K1, ...) {
  n <- length(data)
  scalingConstant <- integrate(function(x){kernel(x, ...)}, -Inf, Inf) |> value()
  function(x) {
    mat <- outer(x, data, FUN = function(x,data){kernel(x-data, ...)})
    val <- rowSums(mat)
    val <- val / (n * scalingConstant)
    return(val)
```

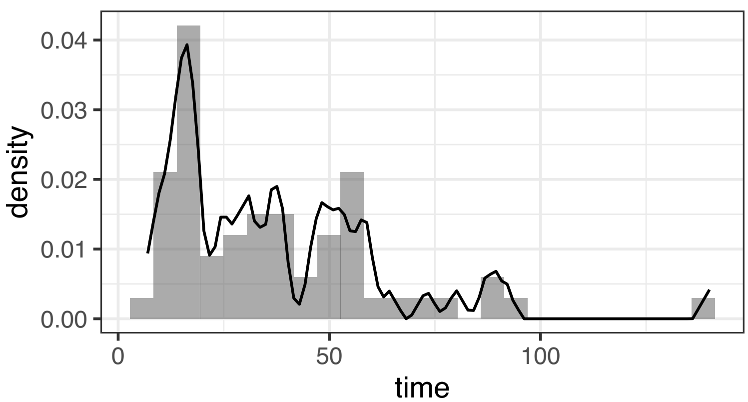
```
  }
}
```

I have selected to fit this same data with a triangle kernel, `K2`, one with increased bandwidth. I have plotted that kernel first to demonstrate how it is nonzero for $|x| < 4$.

```
myDensityCurve <- kde(bballGame$time, kernel = K2)
gf_fun(K2(x) ~ x, xlim=c(-5,5))
```



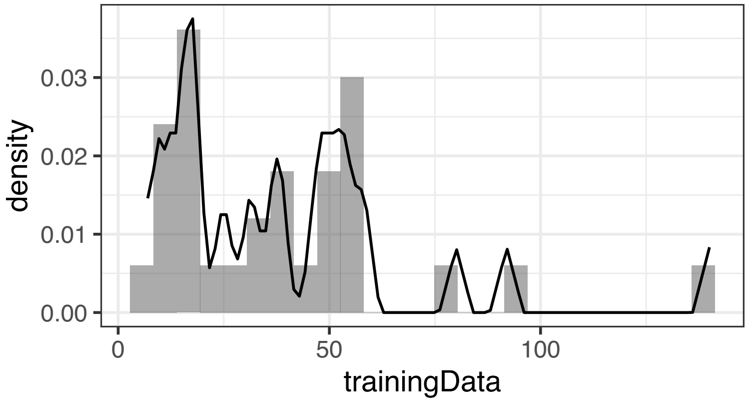

```
gf_dhistogram(~time, data=bballGame) |>
 gf_fun(myDensityCurve(x) ~ x, xlim=c(0,150))
```



Part (d)

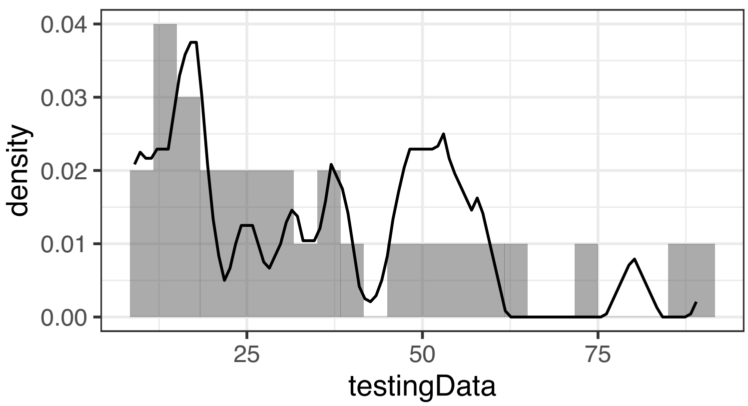
Finally, we select about half of the data for the purpose of training.

```
allIndices = 1:60
trainingIndices <- sample(allIndices, size=30)
testingIndices <- allIndices[-trainingIndices]
trainingData = bballGame[trainingIndices, 1]
testingData = bballGame[testingIndices, 1]
myDensityCurve <- kde(trainingData, kernel = K2)
gf_dhistogram(~trainingData) |>
  gf_fun(myDensityCurve(x) ~ x, xlim=c(0,150))    # just the training data with associated kernel densit
```



The above shows the kernel density function fitted to the training data. It is not surprising that it has peaks in the right places. To see how it works with the remaining (*testing*) data, look at:

```
gf_dhistogram(~testingData) |> gf_fun(myDensityCurve(x) ~ x, xlim=c(0,150))
```



There is evidence here of over-fitting. The peaks that were in sync with the training data do not align with peaks in the testing data.

3.25 $$M_Y(t) = \int_{-\infty}^{\infty} e^{ty} f(y)\,dy = \int_0^{\infty} y e^{ty} e^{-y}\,dy = \int_0^{\infty} y e^{(t-1)y}\,dy$$

$$= \frac{1}{t-1} y e^{(t-1)y}\Big|_0^{\infty} - \frac{1}{t-1}\int_0^{\infty} e^{(t-1)y}\,dy = -\frac{1}{(t-1)^2} e^{(t-1)y}\Big|_0^{\infty}$$

$$= \frac{1}{(t-1)^2}.$$

3.31 $$M_X'(t) = 2e^{2t}(1-t^2)^{-1} + 2te^{2t}(1-t^2)^{-2} \longrightarrow E(X) = M_X'(0) = 2$$

$$M_X''(t) = 4e^{2t}(1-t^2)^{-1} + 8te^{2t}(1-t^2)^{-2} + 2e^{2t}(1-t^2)^{-2} + 8t^2e^{2t}(1-t^2)^{-3}$$

$$\longrightarrow E(X^2) = M_X''(0) = 6$$

Thus, $\mathrm{Var}(X) = 6 - 2^2 = 2.$

3.33 $$M_X'(t) = \frac{18}{(3-t)^3} \longrightarrow E(X) = M_X'(0) = 2/3$$

$$M_X''(t) = \frac{54}{(3-t)^4} \longrightarrow E(X^2) = M_X''(0) = 2/3$$

So, $\mathrm{Var}(X) = \frac{2}{3} - \left(\frac{2}{3}\right)^2 = \frac{2}{9}.$

3.37 (a) Since $X \sim \mathrm{Binom}(n, \pi)$ has MGF $M_X(t) = \left(1 - \pi + \pi e^t\right)^n$,

when $M_X(t) = \left(\frac{1}{2}(e^t+1)\right)^{10}$, $X \sim \mathrm{Binom}(10, 1/2)$.

(b) Since $X \sim \mathrm{Norm}(\mu, \sigma)$ has MGF $M_X(t) = e^{\mu t + \sigma^2 t^2/2}$,

when $M_X(t) = e^{t + t^2/2}$, $X \sim \mathrm{Norm}(1, 1)$.

(c) Since $X \sim \mathrm{Exp}(\lambda)$ has MGF $M_X(t) = \dfrac{1}{1 - t/\lambda}$,

when $M_X(t) = \dfrac{1}{1-2t}$, $X \sim \mathrm{Exp}(1/2)$.

(d) Since $X \sim \mathrm{Gamma}(\alpha, \lambda)$ has MGF $M_X(t) = \dfrac{1}{(1 - t/\lambda)^{\alpha}}$

when $M_X(t) = (1-2t)^{-3}$, $X \sim \mathrm{Gamma}(\alpha = 3, \lambda = 1/2)$, or $\mathrm{Gamma}(\alpha = 3, \beta = 2)$.

3.38 $X \sim \mathrm{Gamma}(\alpha, \lambda)$, so $M_X(t) = \dfrac{1}{(1 - t/\lambda)^{\alpha}}$. Setting $Y = 3X$, we have

$$M_Y(t) = E(e^{tY}) = E(e^{t(3X)}) = E(e^{(3t)X}) = M_X(3t) = \frac{1}{(1 - 3t/\lambda)^{\alpha}}$$

$$\Rightarrow Y \sim \mathrm{Gamma}(\alpha, \lambda/3).$$

3.39 (a) pexp(2) − pexp(0) = 0.865

(b) pexp(1, 2) − pexp(0, 2) = 0.865

(c) $\frac{2}{2\sqrt{3}} \cdot (b-a) \cdot \frac{1}{b-a} = \frac{1}{\sqrt{3}} \doteq 0.5774.$

(d) $E(X) = 1/3$, $Var(X) = 2/63$

diff(pbeta(1/3 + c(-1,1) * sqrt(2/63), 2, 4))

= 0.6522.

3.45 Taking into account that points-scored is always an integer, the normal-quantile plot is quite linear in its appearance, making a normal model appear appropriate. At the top end, there may be the suggestion that he should have had a higher personal-season-best.

3.46 (a) A normal model appears to be reasonable, though there is a slight curve to the normal quantile plot.

(b) The individual plots seem straighter still. There may be some diversion from normality at the extremes.

3.56 Let T_i = lifetime of the i^{th} lightbulb. The cdf of T:

$$F_T(t) = \Pr(T \le t) = \left[\Pr(T_i \le t)\right]^{10} = \left(1 - e^{-t/100}\right)^{10}, \quad \text{for } t \ge 0.$$

3.62 (a) Because $R \sim \text{Norm}(100, 20)$, his obtaining 150 would correspond to a Z-score

$$Z_R = \frac{150-100}{20} = 2.5$$

For $C \sim \text{Norm}(110, 15)$, $$Z_c = \frac{150-110}{15} = 2.667$$

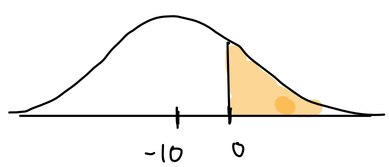
A higher Z-score corresponds to a rarer event. Thus, Ralph should reach scores of 150 (or higher) more often than Claudia.

(b) By normality and independence,

$$R - C \sim \text{Norm}\left(-10, \sqrt{15^2+20^2}\right) = \text{Norm}(-10, 25)$$

$$\Pr(R > C) = 1 - \text{pnorm}(0, -10, 25)$$

$$\doteq 0.345$$



(c) Let $\bar{R}$, $\bar{C}$ be their averages over three games.

$\bar{R} \sim \text{Norm}(100, 20/\sqrt{3})$, $\bar{C} \sim \text{Norm}(110, 15/\sqrt{3})$ and $\bar{R} - \bar{C} \sim \text{Norm}(-10, 25/\sqrt{3})$.

$$\Pr(\bar{R} > \bar{C}) = 1 - \text{pnorm}(0, -10, 25/\sqrt{3}) \doteq 0.244.$$

(d) Let X = # of games won by Ralph. Assuming independence, $X \sim \text{Binom}(3, 0.345)$.

$$\Pr(X \ge 2) = 1 - \text{pbinom}(1, 3, 0.345) = 0.275.$$

3.65 $X \sim \text{Gamma}(\alpha_1, \lambda)$, $Y \sim \text{Gamma}(\alpha_2, \lambda)$, so $M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_1}$, and $M_Y(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_2}$.

X, Y independent means

$$M_{X+Y}(t) = M_X(t)\, M_Y(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_1 + \alpha_2},$$

revealing that $X + Y \sim \text{Gamma}(\alpha_1 + \alpha_2, \lambda)$.

3.66 $X \sim \text{Gamma}(\alpha, \lambda_1)$, $Y \sim \text{Gamma}(\alpha, \lambda_2)$, so $M_X(t) = \left(\frac{\lambda_1}{\lambda_1 - t}\right)^{\alpha}$, and $M_Y(t) = \left(\frac{\lambda_2}{\lambda_2 - t}\right)^{\alpha}$.

X, Y independent means

$$M_{X+Y}(t) = M_X(t)\, M_Y(t) = \left(\frac{\lambda_1 \lambda_2}{(\lambda_1 - t)(\lambda_2 - t)}\right)^{\alpha}.$$

This is not the mgf of a gamma r.v.

C.4 (a) $\text{proj}(\vec{u} \to \vec{v}) = \frac{\langle 1,0\rangle \cdot \langle 1,1\rangle}{\langle 1,1\rangle \cdot \langle 1,1\rangle} \langle 1,1\rangle = \frac{1}{2}\langle 1,1\rangle = \langle \frac{1}{2}, \frac{1}{2}\rangle.$

(b) $\text{proj}(\vec{u} \to \vec{v}) = \frac{\langle 1,0\rangle \cdot \langle 1,-1\rangle}{\langle 1,-1\rangle \cdot \langle 1,-1\rangle} \langle 1,-1\rangle = \frac{1}{2}\langle 1,-1\rangle = \langle \frac{1}{2}, -\frac{1}{2}\rangle.$

(c) $\text{proj}(\vec{u} \to \vec{v}) = \frac{\langle 1,0\rangle \cdot \langle 1,2\rangle}{\langle 1,2\rangle \cdot \langle 1,2\rangle} \langle 1,2\rangle = \frac{1}{5}\langle 1,2\rangle = \langle \frac{1}{5}, \frac{2}{5}\rangle$

(d) $\text{proj}(\vec{u} \to \vec{v}) = \frac{\langle 1,2,3\rangle \cdot \langle 1,1,1\rangle}{\langle 1,1,1\rangle \cdot \langle 1,1,1\rangle} \langle 1,1,1\rangle = \frac{6}{3}\langle 1,1,1\rangle = \langle 2,2,2\rangle.$

(e) $\text{proj}(\vec{u} \to \vec{v}) = \frac{\langle 1,1,1\rangle \cdot \langle 1,2,3\rangle}{\langle 1,2,3\rangle \cdot \langle 1,2,3\rangle} \langle 1,2,3\rangle = \frac{6}{14}\langle 1,2,3\rangle = \langle \frac{3}{7}, \frac{6}{7}, \frac{9}{7}\rangle$

(f) $\text{proj}(\vec{u} \to \vec{v}) = \frac{\langle 1,2,3\rangle \cdot \langle 1,-1,0\rangle}{\langle 1,-1,0\rangle \cdot \langle 1,-1,0\rangle} \langle 1,-1,0\rangle = -\frac{1}{2}\langle 1,-1,0\rangle = \langle -\frac{1}{2}, \frac{1}{2}, 0\rangle$

C.21 This statement is true. To demonstrate it, let $B = (A^T)^{-1}$. Then $I = BA^T$. Taking transposes of both sides and noting $I^T = I$, we have $I = (BA^T)^T = AB^T$, showing that $B^T = A^{-1}$. Transposing again gives $B = (A^{-1})^T$.

C.24 (a) $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = (ad-bc)\, I.$

(b) It is evident that $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, when multiplied by the now-rescaled $\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, is I.

(c) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1} = -\frac{1}{2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}.$