Statistik I

Paul Strimtu 3898312
Jakob Striegel 4351490

Abgabe 3: 22 Januar 2024

1 Fußgängerzone

1.1 a)

Die Situation wird am besten durch eine Geometrische Verteilung beschrieben, da es sich um eine Menge von unabhängigen Bernoulli-Versuchen (Patenschaft abgeschlossen / nicht abgeschlossen) handelt.
Da im Schnitt nur jeder 80. Versuch zum Erfolg führt, ist $p = \frac{1}{80}$.

Nach Vorlesung gilt:

$$E(X) = \frac{1}{p} = \frac{1}{\frac{1}{80}} = 80.$$

1.2 b)

W'keit dafür, dass nach 200 angesprochenen Passanten, noch keine Patenschaft erfolgreich abgeschlossen wurde:

```r
1-pgeom(200-1,1/80)

## [1] 0.08080177
```

1.3 c)

```r
mean(rgeom(1000,1/80)+1)

## [1] 83.13

sum(rgeom(1000,1/80)+1 > 200)/1000

## [1] 0.076
```

Ergebnisse passen grob (da Simulation) zu den Teilergebnissen aus a) und b).

1.4 d)

$P(X \leq k) = 1 - (\frac{79}{80})^k, \quad P(X \leq k) \geq 0,8$

$$\Rightarrow 1 - \left(\frac{79}{80}\right)^k \geq 0,8$$

$$\Leftrightarrow \left(\frac{79}{80}\right)^k \leq 0,2$$

$$\Leftrightarrow k * ln\left(\frac{79}{80}\right) \leq ln(0,2)$$

$$\Leftrightarrow k \geq \frac{ln(0,2)}{ln(\frac{79}{80})} \approx 128$$

```
pgeom(128-1,1/80)

## [1] 0.8001292
```

Es müssen mindestens 128 Passanten befragt werden, damit mit einer W'keit von mindestens 80% eine Patenschaft zustandekommt.

2 Eurostar

2.1 a)

```r
pbinom(894,935,0.95)

## [1] 0.8252784
```

Die W'keit dafür, dass bei 935 verkauften Tickets, alle Passagiere, die erschienen sind, auch einsteigen können, beträgt ca. $82,5\%$.

2.2 b)

```r
pbinom(894,921,0.95)

## [1] 0.9992535

pbinom(894,922,0.95)

## [1] 0.9986704
```

Der Betreiber könnte maximal 921 Tickets verkaufen, sodass mit einer W'keit von $99,9\%$ alle erschienenen Passagiere auch einsteigen können.

2.3 c)

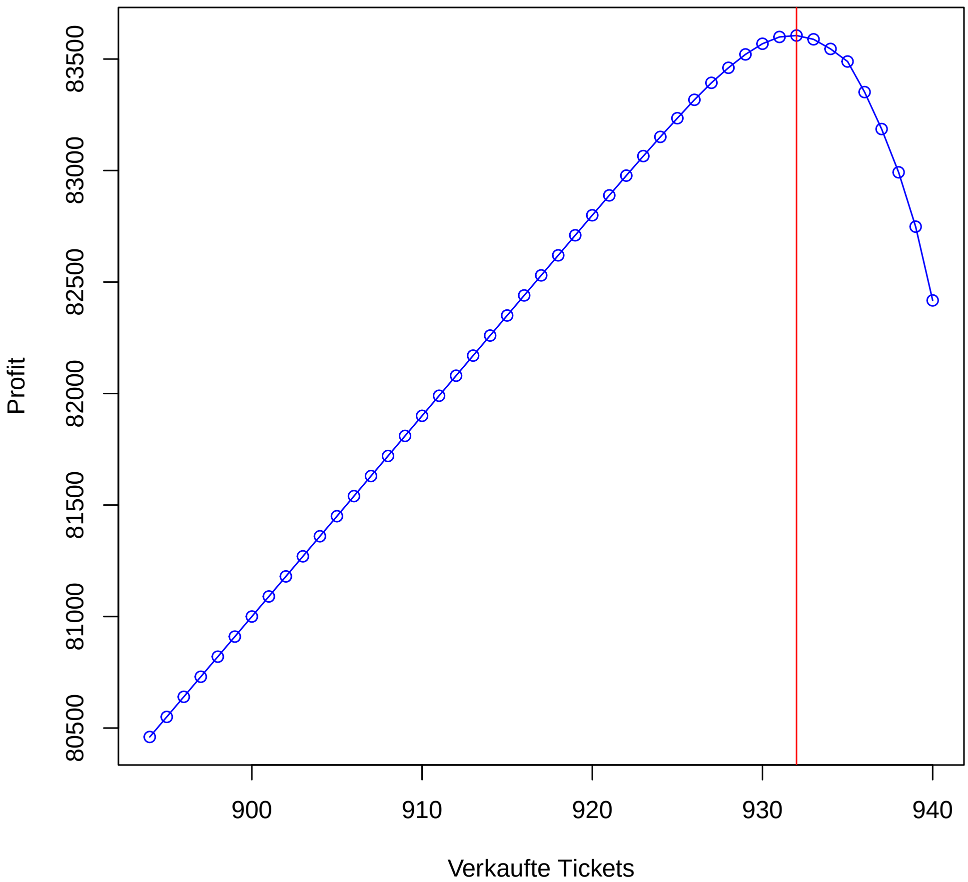
2.3.1 Zufälliges n (sold)

```r
max <- 894
sold <- 930
price <- 90
compensation <- 1000

total <- numeric(1000000)

for (i in 1:1000000) {
  attended <- rbinom(1,sold,0.95)

  if (attended <= max) {
    profit <- sold * price
  } else {
    profit <- sold * price - (attended-max) * compensation
  }

  total[i] <- profit
}

mean(total)

## [1] 83567.72
```

Bei $n = 930$ (sold) ist das obere Ergebnis der durchschnittliche Gewinn beim jeweiligen Simulationsdurchlauf.

2.3.2 Optimales n

```
num <- 894:940
max <- 894
price <- 90
compensation <- 1000

total_in <- numeric(100000)
total_out <- numeric(47)

for (j in 894:940) { #erste Schleife verschiedene #Tickets
  for (i in 1:100000) { #innere Schleife nur 100000 fuer Vereinfachung
    attended <- rbinom(1,j,0.95)

    if (attended <= max) {
      profit <- j * price
    } else {
      profit <- j * price - (attended-max) * compensation
    }

    total_in[i] <- profit
  }
  total_out[j] <- mean(total_in) #speichert fuer 895:940 jeweil
                                 #den mean der jeweiligen profits
}

total_new <- total_out[-c(1:893)] #drop der ersten NA werte fuer total_out

my_data <- data.frame(x = num, y = total_new)

plot(my_data$x, my_data$y, type = "o", col = "blue",
     xlab = "Verkaufte Tickets", ylab = "Profit")
abline(v=932,col="red")
```




```
my_data[my_data$y == max(my_data$y),]

##       x       y
## 39 932 83605.4
```

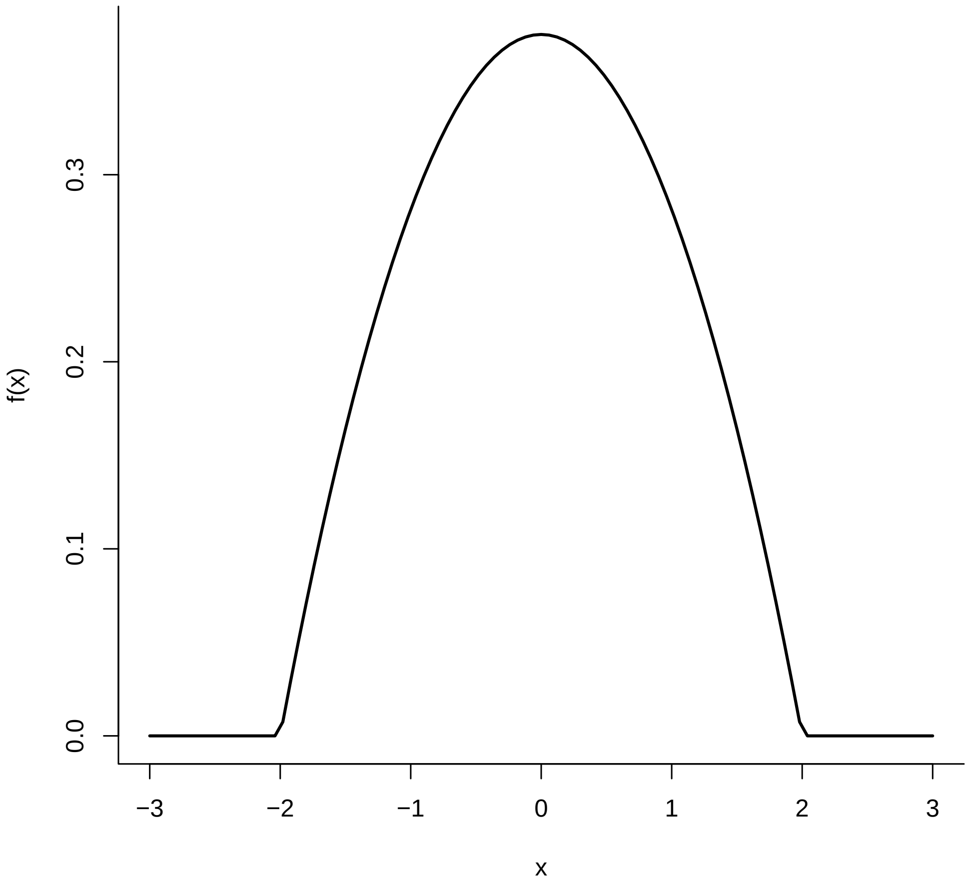
Das kostenoptimale n liegt in diesem Fall bei 932. Der Betreiber erzielt also bei 932 verkauften Tickets den höchsten Gewinn, wobei der Gewinn aufgrund der zufälligen Simulation leicht variieren kann.

3 Dichtefunktion & Verteilungsfunktion

3.1 a)

$$f(x) = \begin{cases} \frac{3}{8}(1 - \frac{x^2}{4}) & \text{für } -2 \leq x \leq 2, \\ 0 & \text{sonst.} \end{cases}$$

```
f <- function(x){
  fx <- ifelse(abs(x)>2,0,3/8*(1-(x^2)/4))
  return(fx)
}

curve(f(x),-3,3,lwd=2,bty="l")
```



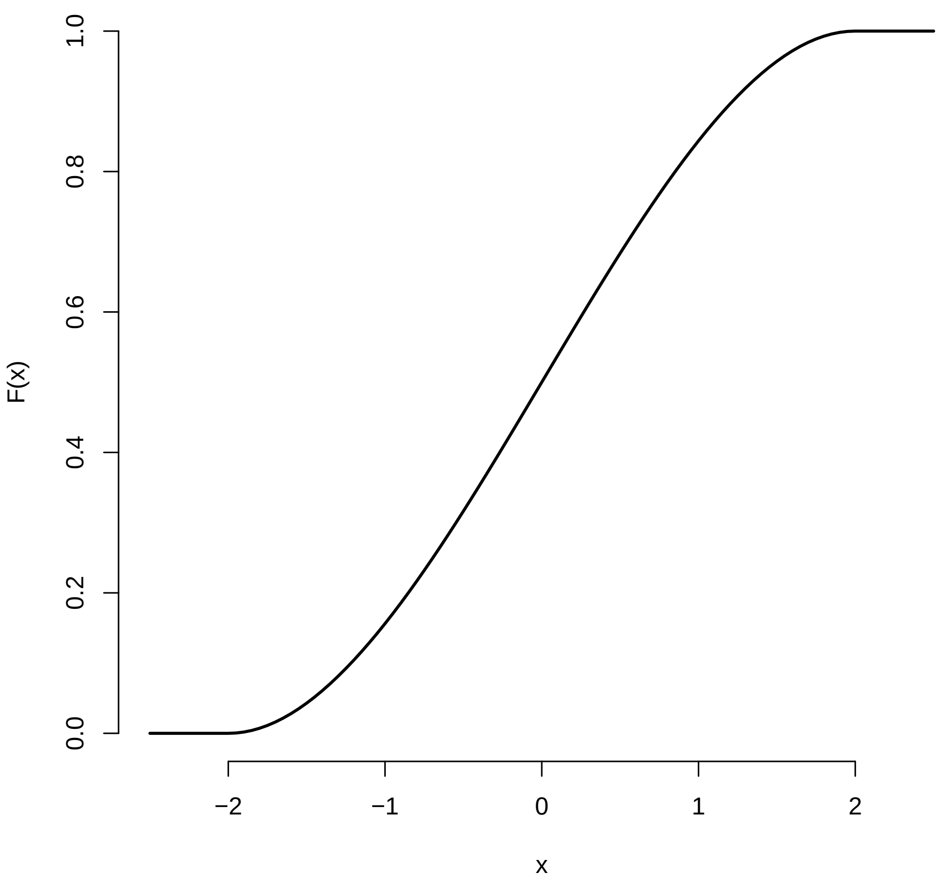
3.2 b)

$$
\begin{aligned}
F(x) &= \int \frac{3}{8}\left(1-\frac{x^2}{4}\right)dx \\
&= \frac{3}{8}\int\left(1-\frac{x^2}{4}\right)dx \\
&= \frac{3}{8}\left(x-\frac{x^3}{12}\right)
\end{aligned}
$$

$$
\Rightarrow F(x) = \begin{cases} 0 & \text{für } x < -2, \\ \frac{3}{8}\left(x-\frac{x^3}{12}\right)+\frac{1}{2} & \text{für } -2 \leq x \leq 2, \\ 1 & \text{für} x > 2. \end{cases}
$$

Die $+\frac{1}{2}$ ist als Korrektur für die Darstellung der Verteilungsfunktion da, sodass die y-Skala Werte von 0 bis 1 und nicht -0.5 bis 0.5 anzeigt.

```
F <- function(x){
  Fx <- ifelse(x< -2,0,ifelse(x>2,1,3/8*(x-(x^3)/12)+1/2))
  return(Fx)
}

curve(F(x),-2.5,2.5,lwd=2,bty="n")
```



3.3 c)

3.3.1 $P(0 < X < 1)$ mit R-Befehl

```
integrate(f,0,1)$value

## [1] 0.34375
```

3.3.2 $P(0 < X < 1)$ mit Verteilungsfunktion

```
F(1)-F(0)

## [1] 0.34375
```

4 Erwartungswert, Varianz & Quantil

4.1 a)

$$
\begin{aligned}
E(x) &= \int_{-2}^{2} xf(x)dx \\
&= \int_{-2}^{2} \left(\frac{3}{8}x\left(1-\frac{x^2}{4}\right)\right) dx \\
&= \frac{3}{8}\int_{-2}^{2} x - \frac{x^3}{4} dx \\
&= \frac{3}{8}\left[\frac{x^2}{2} - \frac{x^4}{16}\right]_{-2}^{2} \\
&= \frac{3}{8}((2-1)-(2-1)) \\
&= \frac{3}{8}(1-1) \\
&= 0
\end{aligned}
$$

4.2 b)

$$
\begin{aligned}
\int (x-E(x))^2 f(x)dx &= \int \frac{3}{8}x^2(1-\frac{x^2}{4})dx \\
&= \int \left(\frac{3x^2}{8} - \frac{3x^2}{32}\right) dx \\
&= \frac{3}{8}\int x^2 dx - \frac{3}{32}\int x^4 dx \\
&= (\frac{3}{8} * \frac{x^3}{3}) - (\frac{3}{32} * \frac{x^5}{5}) \\
&= \frac{x^3}{8} - \frac{3x^5}{160}
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow Var(X) &= \int_{-2}^{2} (x-E(x))^2 f(x)dx \\
&= \left[\frac{x^3}{8} - \frac{3x^5}{160}\right]_{-2}^{2} \\
&= \left(1-\frac{96}{160}\right) - \left(-1+\frac{96}{160}\right) \\
&= \frac{64}{160} + \frac{64}{160} \\
&= \frac{128}{160} \\
&= \frac{4}{5}
\end{aligned}
$$

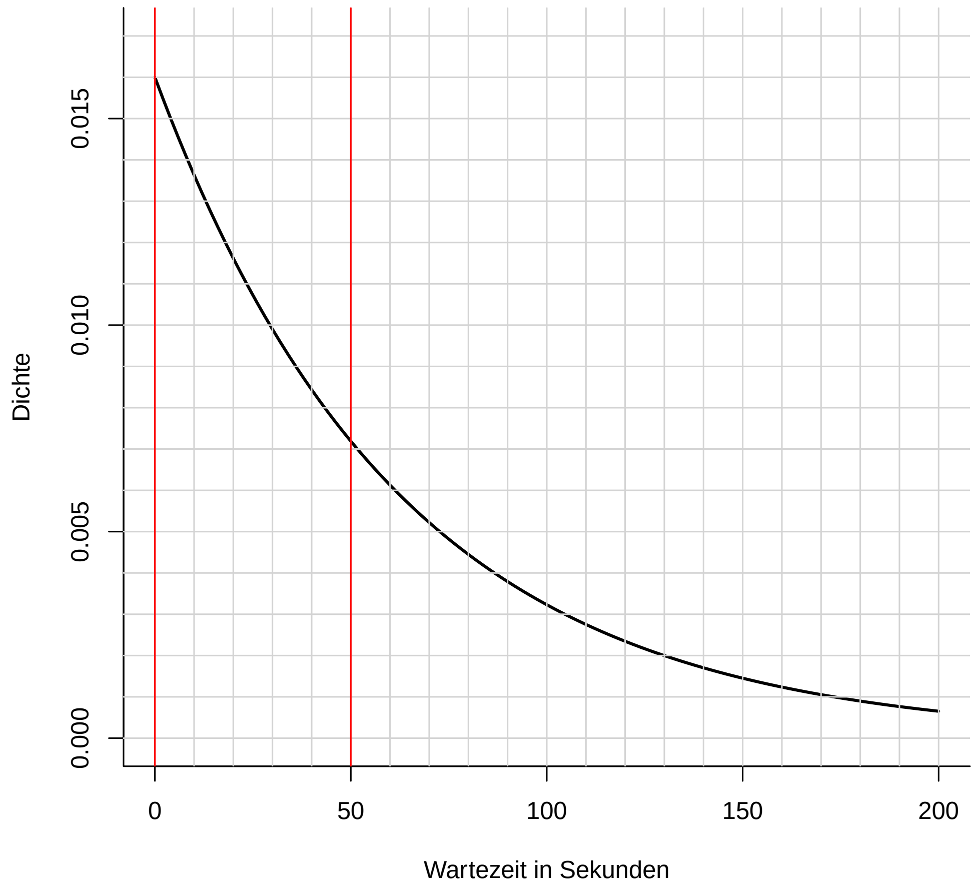
4.3 c)

```
F(0.851437)

## [1] 0.8
```

5 Kaffee bei Hasan

5.1 Abschätung des Flächeninhaltes

```
curve(dexp(x,0.016),0,200,lwd=2,bty="l",ylim=c(0,0.017)
  ,xlab = "Wartezeit in Sekunden", ylab = "Dichte")

abline(v=seq(0,200,by=10),col="lightgrey")
abline(h=seq(0,0.02,by=0.001),col="lightgrey")
abline(v=c(0,50),col="red")
```



$\square \widehat{=} 0,01 = 0,001 * 10$
Bei ca. 55 Kästchen beträgt die Fläche ungefähr 0.55 FE.

5.2 Integration

5.2.1 Numerisch

```
k <- function(x){
  kx <- 0.016*exp(-0.016*x)
  return(kx)
}

integrate(k,0,50)$value

## [1] 0.550671
```

5.2.2 Exakt

$$
\begin{aligned}
\int_0^{50} k(x)dx &= \int_0^{50} \frac{2}{125} e^{(\frac{-2x}{125})} dx \qquad\qquad u = \frac{2x}{125}, \quad \frac{du}{dx} = \frac{2}{125}, \quad dx = \frac{125du}{2} \\
&= \int_0^{50} e^{-u} du \\
&= \left[-e^{\frac{-2x}{125}} \right]_0^{50} \\
&= -e^{-\frac{100}{125}} + e^0 \\
&= 1 - e^{-\frac{4}{5}}
\end{aligned}
$$

5.3 R-Befehl

```
pexp(50,0.016)

## [1] 0.550671
```