Segunda avaliação
Prof. Thiago Gentil Ramires - Probabilidade e Estatística

Nome: Deivid da Silva Galvão

RA: 2408740

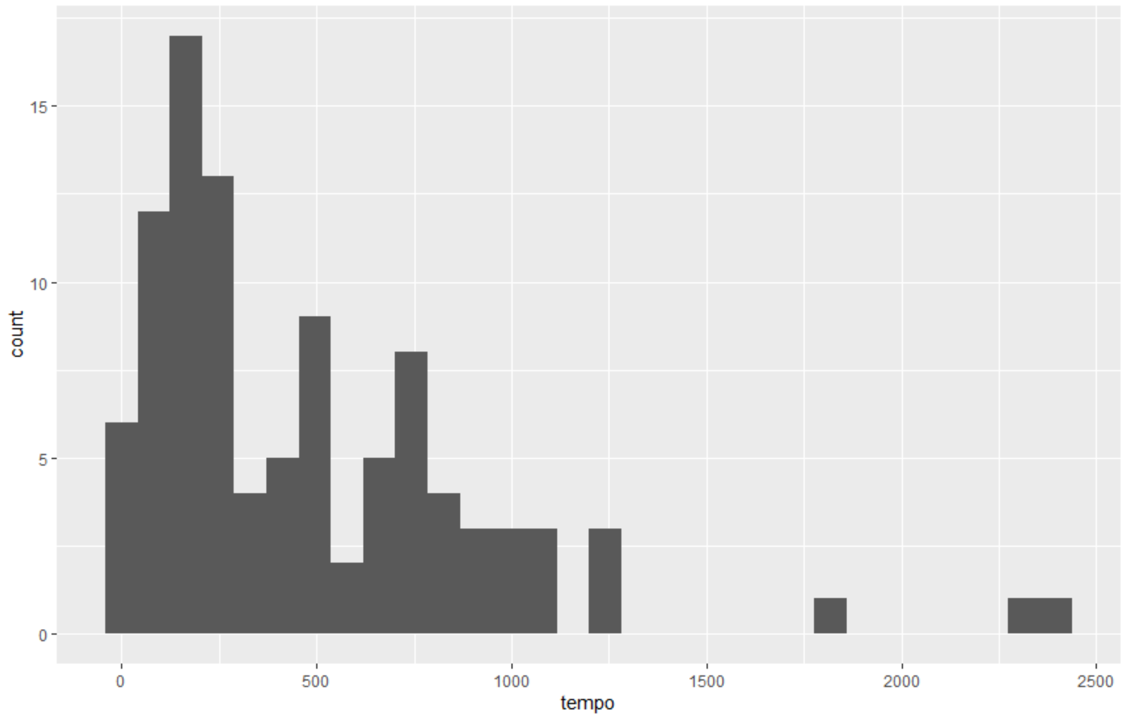
Disciplina matriculado(a): Probabilidade e Estatística – Eng. Comp.

Para todas as questões abaixo, **interprete os resultados e apresente os códigos e gráficos, quando necessário**. As bases de dados estão em anexo do Google Classroom, já salvas em CSV, com o separador decimal em Inglês, ou seja, as decimais estão separadas por ponto. **Cada questão vale 0,5 pontos.**

Questão 1) O tempo de duração (em horas) de 100 vigas metálicas, após teste de força, está apresentado no arquivo ex1.csv. Determine:

a) Qual o melhor modelo de probabilidade que representa tais tempos?

Resposta: Exponencial, pois o grafico gerado apresenta uma tendencia de função exponencial.



Códigos:

```
install.packages('ggplot2')
library(ggplot2)
dados=read.csv('ex1.csv')
ggplot(dados,aes(tempo))+geom_histogram()
```

b) Determine a probabilidade de uma viga durar menos de 300 horas.

Resposta: 0.464611

Códigos:

```
pexp(300,1/mean(dados$tempo))
```

c) Determine a probabilidade de uma viga durar entre 200 e 400 horas.

Resposta: 0.8839544

probabilidade de ser maior que 200 + probabilidade de ser melhor que 400 - probabilidade de ser menor que 200

Códigos:

```
a = 1-pexp(200,1/mean(dados$tempo))
b = pexp(400,1/mean(dados$tempo))
c = -pexp(200,1/mean(dados$tempo))
sum(a,b,c)
```

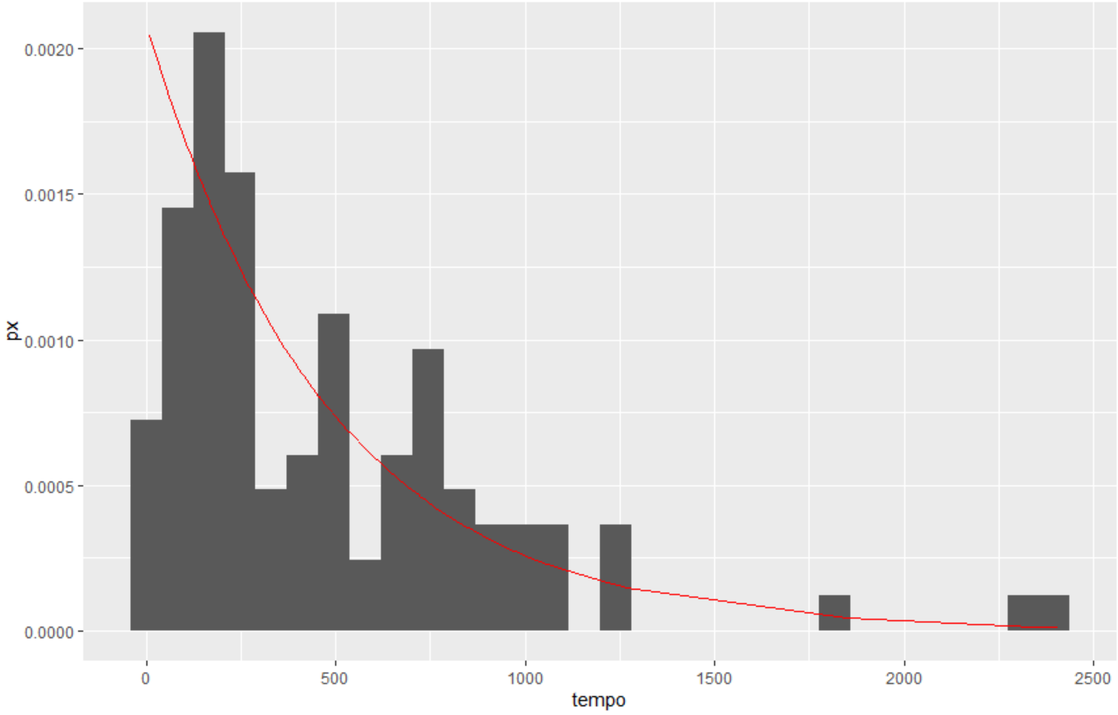
d) Serão descartadas 70% das vigas com menor tempo de duração. Determine o tempo ideal de descarte, ou seja, o tempo que limita o descarte das vigas.

Resposta:

Códigos:

e) Apresente graficamente o histograma dos dados juntamente com o modelo escolhido (ajustado) na resposta A).

Resposta: Com a linha de ajuste é possível confirmar que o modelo escolhido é realmente adequado para esses dados.



Códigos:

```
dados$px=dexp(dados$tempo,1/mean(dados$tempo))

ggplot(dados,aes(tempo,px))+geom_histogram(aes(y=
..density..))+geom_line(col='red')
```

Questão 2) Uma empresa está interessada em estudar o comportamento dos produtos eletrônicos produzidos por ela. Em um teste com 20 desses produtos produzidos, 6 apresentaram defeitos. Em uma semana foram produzidos 100 novos produtos. Determine:

a) A probabilidade de mais de 75 não apresentarem defeito?

Resposta:

$p = 6/20 = 0,3$

Códigos:

b) A probabilidade de exatamente 70 não apresentarem defeito?

Resposta:

Códigos:

c) A probabilidade de menos de 20 produtos apresentarem defeito?

Resposta: 0.9911128

Códigos:

```
1-pbinom(19,100,30/100)
```

d) Se no próximo mês forem produzidos 4000 produtos, quantos irão falhar em média?

Resposta: 1200

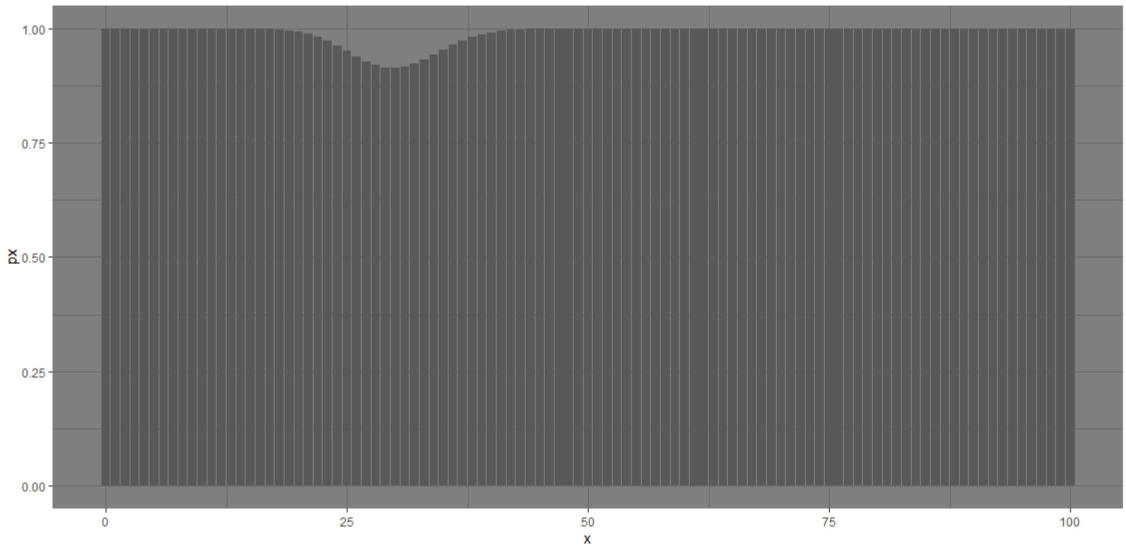
numero de produtos * a probabilidade de falhar

Códigos:

```
4000*0.3
```

e) Apresente graficamente todas as probabilidades de não apresentar falha, considerando os n=100 produtos.

Resposta:



Códigos:

```
x=0:100
px=1-dbinom(0:100,100,0.3)
dados = data.frame(x,px)
#plotando todas as probabilidades
ggplot(dados,aes(x,px))+ geom_col()+
```

```
scale_fill_brewer(palette = "Set1") + theme_dark()
```

Questão 3) Um algoritmo de detecção de anomalias capta, em média, 20 erros por hora. Determine:

a) A probabilidade de detectar 15 erros em uma hora?

Resposta: 0.05164885,

funcao do modelo poisson com media de 20 por hora.

Códigos:

```
dpois(15, 20)
```

b) A probabilidade de detectar entre 20 e 30 erros em uma hora?

Resposta: 0.449251

probabilidade de 30 erros em 1h - probabilidade de ser menor que 20 em 1h

Códigos:

```
a=dpois(30,20)

b=1-ppois(20,20)

sum(a,b)
```

c) A probabilidade de detectar mais de 449 erros em um dia?

Resposta: 0.08081902

funcao do modelo poisson com media de 20 * 24(1 dia).

Códigos:

```
ppois(449, 20*24)
```

d) Um novo algoritmo foi testado, sendo que a média de detecção em uma hora foi de 30 erros. Esse algoritmo será adquirido pela empresa caso detecte na próxima hora mais de 34 erros. Determine a probabilidade da compra ser realizada.

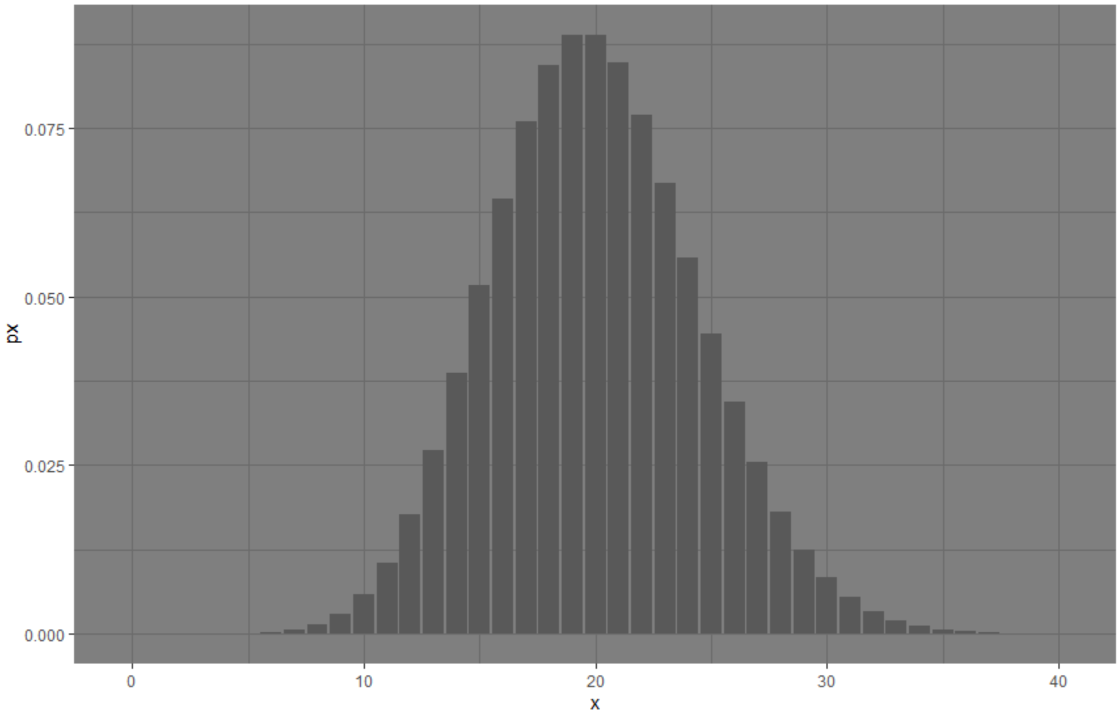
Resposta: 0.2026917

Códigos:

```
1-ppois(34,30)
```

e) Apresente o gráfico das probabilidades de detecção de erros do algoritmo, considerando o algoritmo com média de 20 erros por hora.

Resposta: A saida foi um grafico de colunas com as probabilidades de erros do algoritmo.(com media de 20 erros por hora).



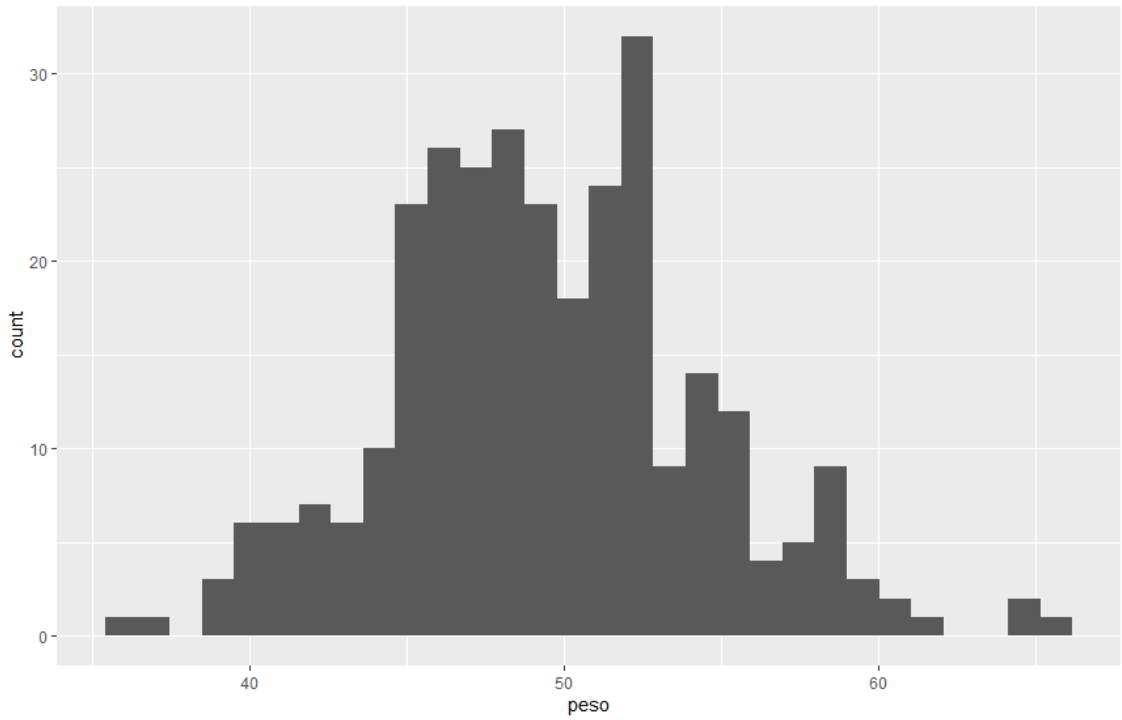
Códigos:

```
x=0:40
px=dpois(x,20)
dados2=data.frame(x,px)
ggplot(dados2,aes(x,px))+geom_col()+theme_dark()
```

Questão 4) O peso de um produto, em Kg, foi determinado em uma amostra de tamanho 300 (Ver anexo ex4.csv). Determine:

a) Qual o melhor modelo de probabilidade que representa tais pesos?

Resposta: Por conta do grafico gerado, podemos concluir que o melhor modelo é o modelo normal.



Códigos:

dados2=read.csv('ex4.csv')

ggplot(dados2,aes(peso))+geom_histogram()

b) Determine a probabilidade de um produto não pesar entre 45 e 55kg.

Resposta: 0.3137118

Códigos:

```
1-(pnorm(55,mean(dados2$peso),sd(dados2$peso))-pnorm(45,mean(dados2$peso),sd(dados2$peso)))
```

c) Determine a probabilidade de um peso ter exatamente 50Kg.

Resposta: Zero , pois no modelo normal a probabilidade de ser exatamente um valor em especifico é sempre 0.

d) Os produtos serão classificados entre: leves (30% mais baixos), médios (pesos entre 30 a 70%) e pesados (os 30% maiores). Determine os limites dos pesos para realizar essa classificação.

Resposta: leves = 46.90295

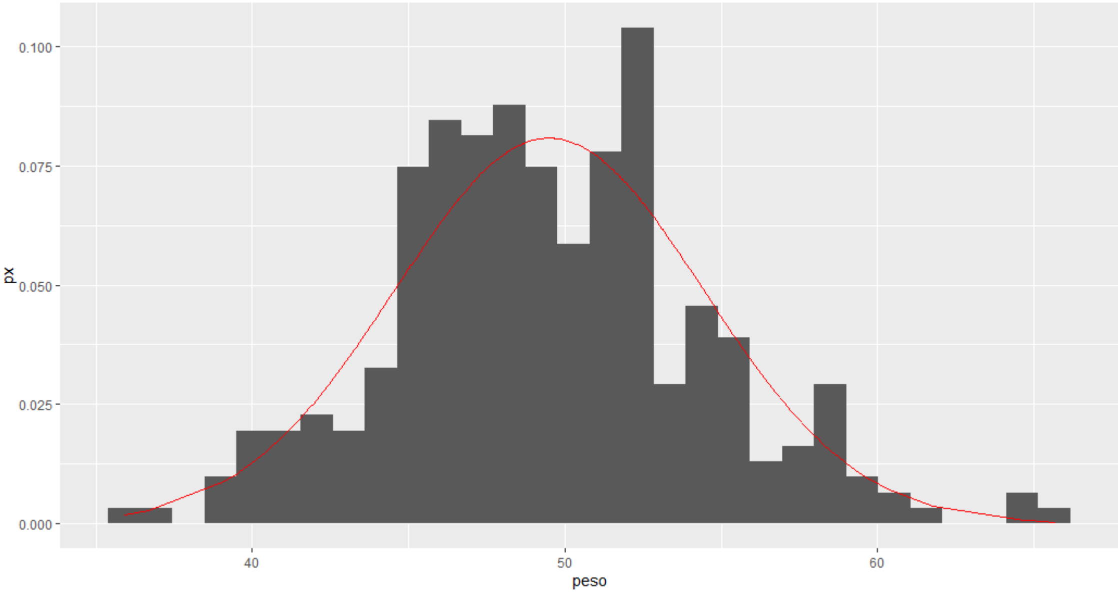
Códigos:

```
#leves
qnorm(0.3,mean(dados2$peso),sd(dados2$peso))
#medios
#pesados
```

e) Apresente graficamente o histograma dos dados juntamente com o modelo escolhido (ajustado) na resposta A).

Resposta: Com a linha de ajuste é possível confirmar que o modelo escolhido é realmente adequado para esses dados.

Segunda avaliação
Prof. Thiago Gentil Ramires - Probabilidade e Estatística



Códigos:

dados2$px=dnorm(dados2$peso,mean(dados2$peso),sd(dados2$peso))

ggplot(dados2,aes(peso,px))+geom_histogram(aes(y= ..density..))+geom_line(col='red')