



UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE MATEMÁTICA, ESTATÍSTICA
E COMPUTAÇÃO CIENTÍFICA



GRUPO 7

Lucas Magri Bueno, **RA:** 201996

Lucca Ferreira Paiva, **RA:** 240229

Maicon Gabriel de Oliveira, **RA:** 221329

Susane Amabila Freires dos Santos, **RA:** 193978

Trabalho ME323 – Introdução aos Modelos Probabilísticos

CAMPINAS
2021

Parte 1

i) Sexo - Qualitativa Nominal, Educação - Qualitativa Ordinal, Cargo - Qualitativa Nominal, Salário - Quantitativa Contínua e Indicador de incremento salarial - Quantitativa Discreta

ii) Distribuição de frequências para cada variável.

VARIÁVEL Sexo:

Companhia A	
Sexo	Frequência
Feminino	126
Masculino	173
Total	299

Companhia B	
Sexo	Frequência
Feminino	118
Masculino	129
Total	247

VARIÁVEL Cargo:

Companhia A	
Cargo	Frequência
Gerencial	46
Outro	253
Total	299

Companhia B	
Cargo	Frequência
Gerencial	30
Outro	217
Total	247

VARIÁVEL Educação:

Companhia A	
Educação	Frequência
Primária	129
Secundária	128
Terciária	42

Companhia B	
Educação	Frequência
Primária	112
Secundária	116
Terciária	19

VARIÁVEL Salário:

Companhia A	
Salário	Frequência
(0, 4]	45
(4, 8]	113
(8, 12]	90
(12, 16]	16
(16, 20]	30
(20, 24]	5

Companhia B	
Salário	Frequência
(0, 4]	76
(4, 8]	131
(8, 12]	24
(12, 16]	15
(16, 20]	1
(20, 24]	0

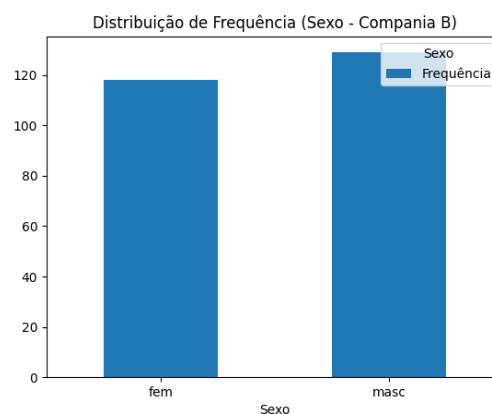
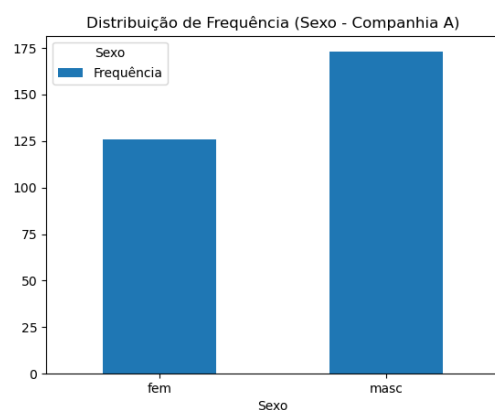
VARIÁVEL Incremento Salarial:

Companhia A	
Incremento Salarial	Frequência
0	639
1	553

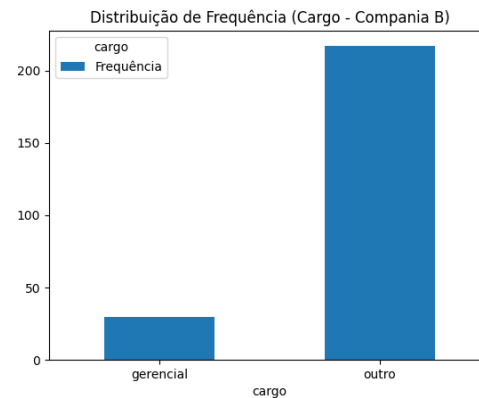
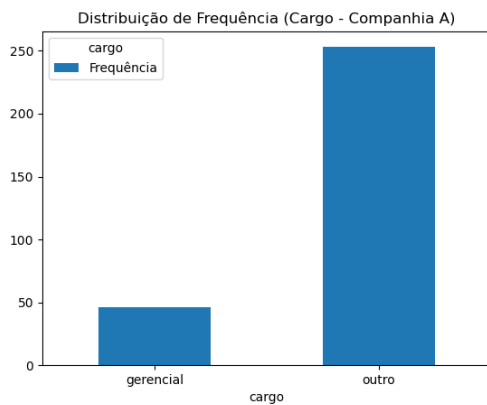
Companhia B	
Incremento Salarial	Frequência
0	540
1	500

iii) Gráfico de distribuição de frequências para cada variável.

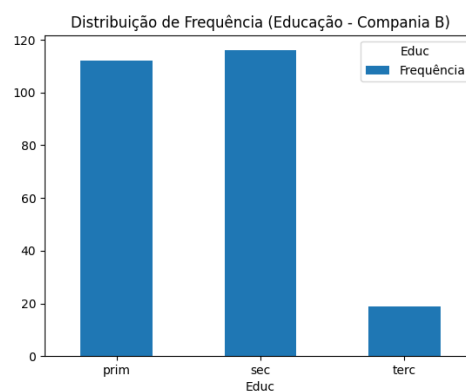
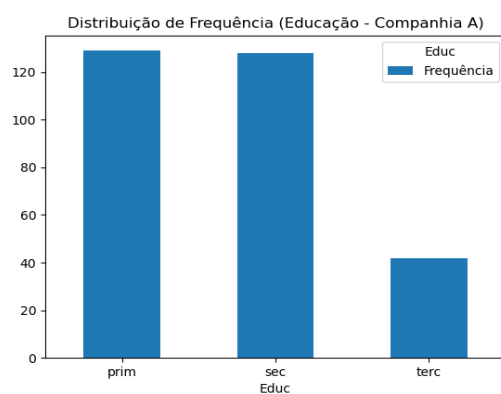
VARIÁVEL Sexo:



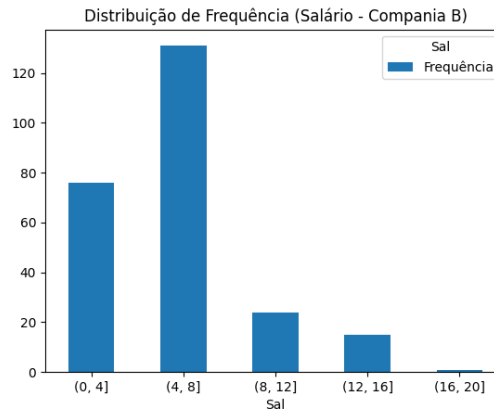
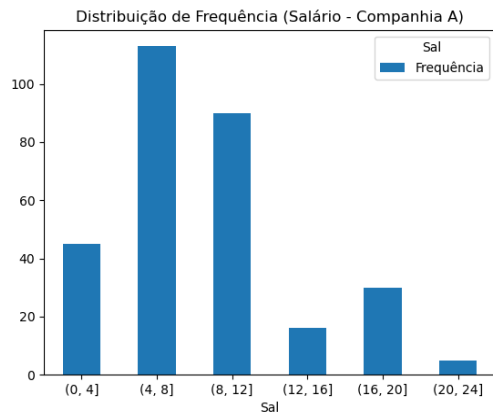
VARIÁVEL Cargo:



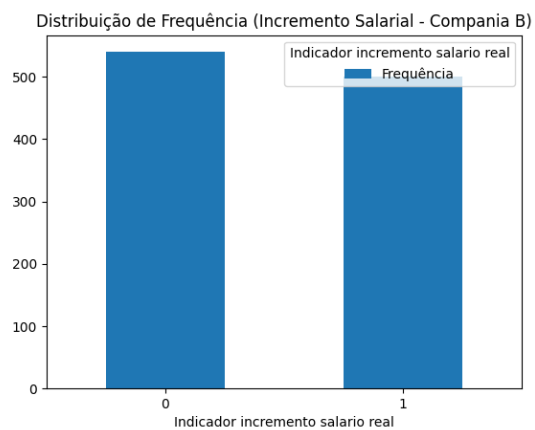
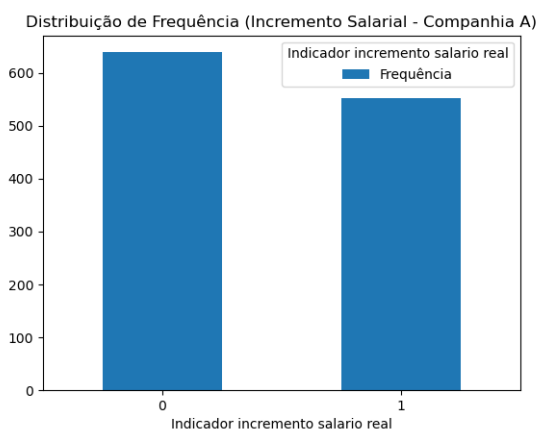
VARIÁVEL Educação:



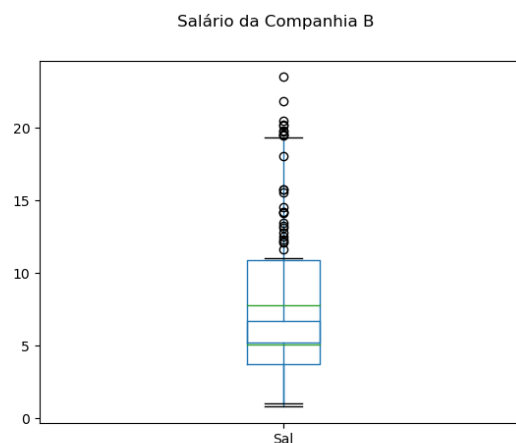
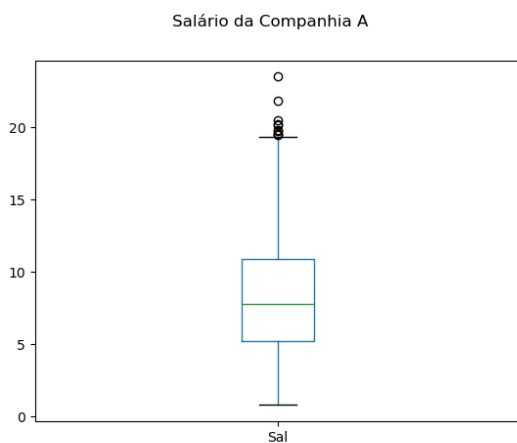
VARIÁVEL Salário:



VARIÁVEL Incremento salarial:



iv) Boxplot para variável salário.



vi) A primeira vista percebe-se uma presença masculina maior em ambas as companhias, a maior parte de seus funcionários não se encontram em cargos gerenciais, os níveis de educação predominantes são primário e a maior frequência de salários se concentram entre 4 a 8 salários.

Parte 2

i) Dependência entre Gênero e Grau de Escolaridade

Companhia A (%)			
	Masculino	Feminino	Total
Primária	34	56	43
Secundária	49	34	43
Terciária	17	10	14
Total	100	100	100

Distribuição Conjunta de Gênero segundo Grau de Escolaridade na companhia A (em porcentagem)

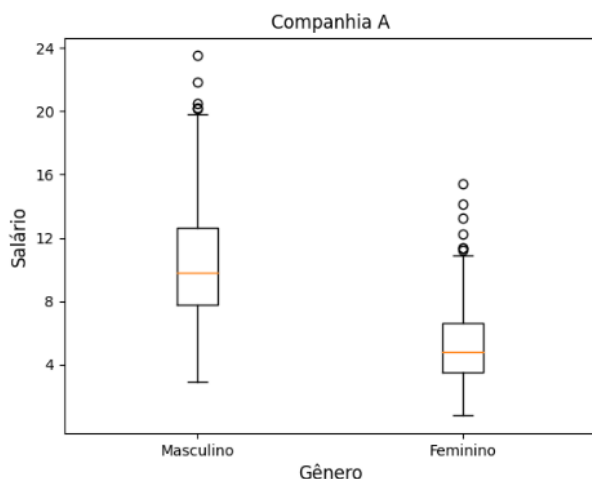
Companhia B (%)			
	Masculino	Feminino	Total
Primária	44	47	45
Secundária	49	45	47
Terciária	7	8	8
Total	100	100	100

Distribuição Conjunta de Gênero segundo Grau de Escolaridade na companhia B (em porcentagem)

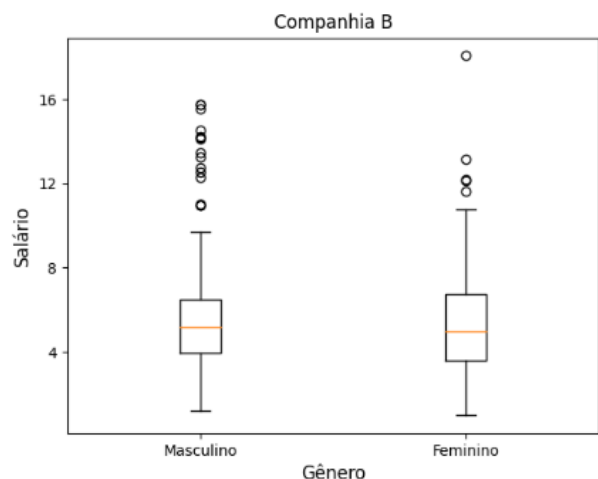
Considerando 2 pontos como margem de erro, para companhia A, gênero e grau de escolaridade se apresentam como fenômenos associados e logo dependentes, pois através da tabela Distribuição Conjunta de Gênero segundo Grau de Escolaridade na companhia A (em porcentagem) observa-se certa discrepância entre o total esperado de pessoas por grau de educação e o efetivo obtido para cada sexo. Para confirmar tal discrepância o grupo calculou o Qui-quadrado de Pearson através da tabela Distribuição Conjunta de Gênero segundo Grau de Escolaridade na companhia A e obteve o valor aproximado de 10,32, o que devido a distância do valor 0, confirma existir certa associação entre as variáveis analisadas. Analisando tais variáveis constata-se que nessa companhia a maioria das mulheres tem educação primária (56%), a maioria dos homens possui educação secundária (49%), há mais homens (17%) do que mulheres (10%) com educação terciária e consequentemente concluiu-se que nessa empresa homens tendem a ter maior grau de instrução do que mulheres.

Para a companhia B, gênero e grau de escolaridade não se apresentam como fenômenos associados e logo não possuem dependência, pois os valores obtidos entre as variáveis gênero por grau de escolaridade se apresentam todos contidos dentro da margem de erro do total esperado. Para confirmar tal tendência o grupo calculou o Qui-quadrado de Pearson através da tabela Distribuição Conjunta de Gênero segundo Grau de Escolaridade na companhia B e obteve o valor 0, comprovando não existir associação significativa entre as variáveis.

ii) Dependência entre Gênero e Renda



Box Plot de Distribuição Conjunta de Gênero segundo Faixa Salarial na companhia A

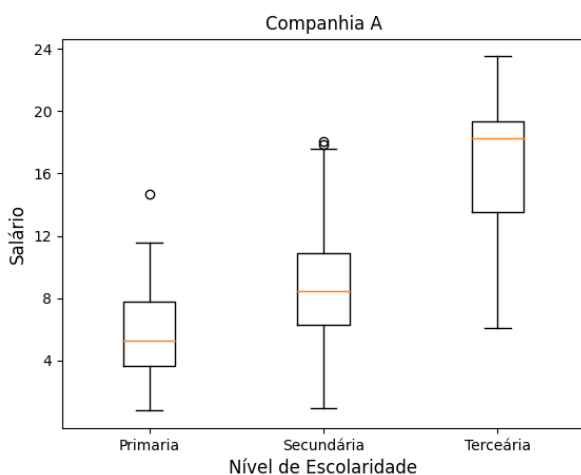


Box Plot de Distribuição Conjunta de Gênero segundo Faixa Salarial na companhia B

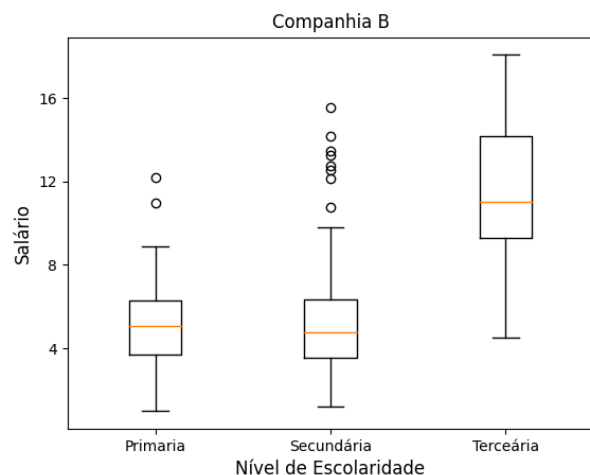
Para a companhia A, através da análise do Box Plot de Distribuição Conjunta de Gênero segundo Faixa Salarial na companhia A constatou-se haver dependência entre as variáveis.

Utilizando a média das variâncias $\overline{var(S)} = \frac{\sum_{i=1}^k n_i var_i(s)}{\sum_{i=1}^k n_i} = 15,22$ e logo $R^2 = 1 - \frac{var(s)}{\overline{var(s)}} = 0,34$ comprovou-se uma dependência de 34% entre as variáveis. Analisando tais dados observou-se haver uma maior amplitude salarial para o gênero masculino, abrangendo de cerca de 3 até 20 salários, assim como a grande maioria destes concentrados entre 8 a 13 salários. Para o gênero feminino, fora os desvios, observa-se uma amplitude de cerca de 1 até 11 salários, onde grande maioria se concentra entre 3 a 6 salários. Constatando-se assim uma tendência de que nessa companhia homens tendem a ter maior remuneração do que mulheres.

Para a companhia B, através da análise do Box Plot de Distribuição Conjunta de Gênero segundo Faixa Salarial na companhia B constatou-se não haver dependência entre as variáveis. Pois ambas as amplitudes de salário e mediana se encontram aproximadamente na mesma região. Utilizando-se a média das variâncias $\overline{var(S)} = 9,37$ e logo $R^2 = 0$, onde pela inexistência de associação, comprova-se não haver dependência entre as variáveis analisadas.



Box Plot de Distribuição Conjunta de Grau de Escolaridade segundo Renda na companhia A



Box Plot de Distribuição Conjunta de Grau de Escolaridade segundo Renda na companhia B

iii) Dependência entre Grau de Escolaridade e Renda

Para ambas as companhias A e B, através da análise de seus respectivos gráficos Box Plot de Distribuição Conjunta de Grau de Escolaridade segundo Renda na companhia A e Box Plot de Distribuição Conjunta de Grau de Escolaridade segundo Renda na companhia B constatou-se haver certa dependência entre as variáveis analisadas. Para companhia A, utilizando a média das variâncias $\overline{var(S)} = \frac{\sum_{i=1}^k n_i var_i(s)}{\sum_{i=1}^k n_i} = 10,35$ e logo $R^2 = 1 - \frac{var(s)}{\overline{var(s)}} = 0,55$ quantificou-se uma associação de 55% entre as variáveis e, para companhia B, também utilizando a média das variâncias $\overline{var(S)} = 6,62$ e logo $R^2 = 0,30$ quantificou-se uma associação de 30% entre as variáveis. Para ambos os casos observa-se que a renda obtida tende a crescer proporcionalmente ao grau de escolaridade de cada funcionário, havendo assim não só maior amplitude para aqueles com escolaridade secundária comparados a primária como também concentração de salários superando todas as demais para funcionários com escolaridade terciária.

iv) Dependência entre Gênero e Cargo Ocupado

Companhia A (%)			
	Masculino	Feminino	Total
Gerencial	20	10	15
Outro	80	90	85
Total	100	100	100

Distribuição Conjunta de Gênero segundo Cargo Ocupado na companhia A (em porcentagem)

Companhia B (%)			
	Masculino	Feminino	Total
Gerencial	10	14	12
Outro	90	86	88
Total	100	100	100

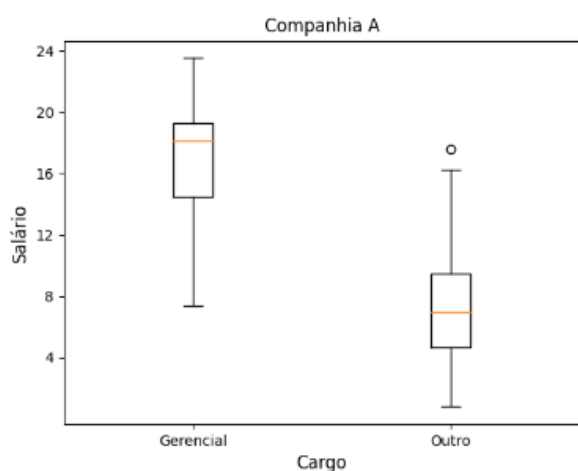
Distribuição Conjunta de Gênero segundo Cargo Ocupado na companhia B (em porcentagem)

Considerando 2 pontos como margem de erro, para companhia A, gênero e cargo ocupado se apresentam como fenômenos associados e logo dependentes, pois através da tabela Distribuição Conjunta de Gênero segundo Cargo Ocupado na companhia A (em porcentagem) observa-se certa discrepância entre o total esperado de pessoas por tipo de cargo ocupado e o efetivo obtido para cada sexo, pois esperou-se um total de 15% para cargos gerenciais e obteve-se 20% sendo do sexo masculino e 10% do sexo feminino. Para confirmar tal tendência, o grupo calculou o Qui-quadrado de Pearson e obteve o valor aproximado de 3,92, confirmando existir uma tímida associação entre as variáveis. Analisando tal dependência percebe-se que o número de mulheres em cargos não gerenciais (90%) é ligeiramente maior do que o número de homens em cargos não gerenciais (80%), levando assim a uma possível, porém fraca tendência de que nessa companhia mulheres são mais propensas a alcançar cargos gerenciais.

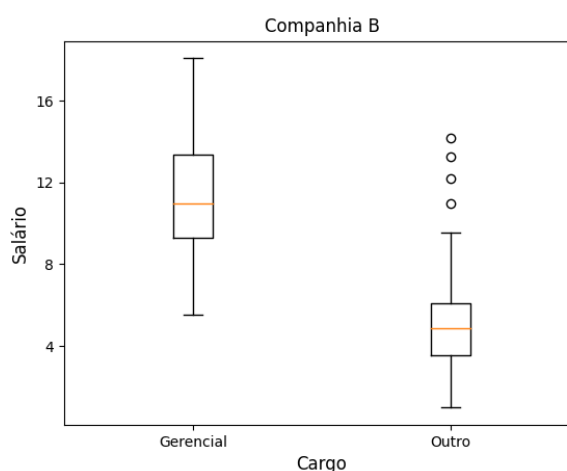
Para companhia B, segundo a análise da tabela Distribuição Conjunta de Gênero segundo Cargo Ocupado na companhia B (em porcentagem) gênero e cargo ocupado não se apresentam como fenômenos associados e logo não possuem dependência,

pois os valores obtidos entre as variáveis gênero e cargo ocupado se apresentam todos contidos dentro da margem de erro do total esperado. Para confirmar tal tendência o grupo calculou o Qui-quadrado de Pearson através da tabela Distribuição Conjunta de Gênero segundo Cargo Ocupado na companhia B e obteve o valor aproximado de 0,76, o que devido à proximidade de 0, comprova a não existência de dependência entre as variáveis analisadas.

v) Dependência entre Cargo Ocupado e Renda



Box Plot de Distribuição Conjunta de Cargo Ocupado segundo Renda na companhia A



Box Plot de Distribuição Conjunta de Cargo Ocupado segundo Renda na companhia B

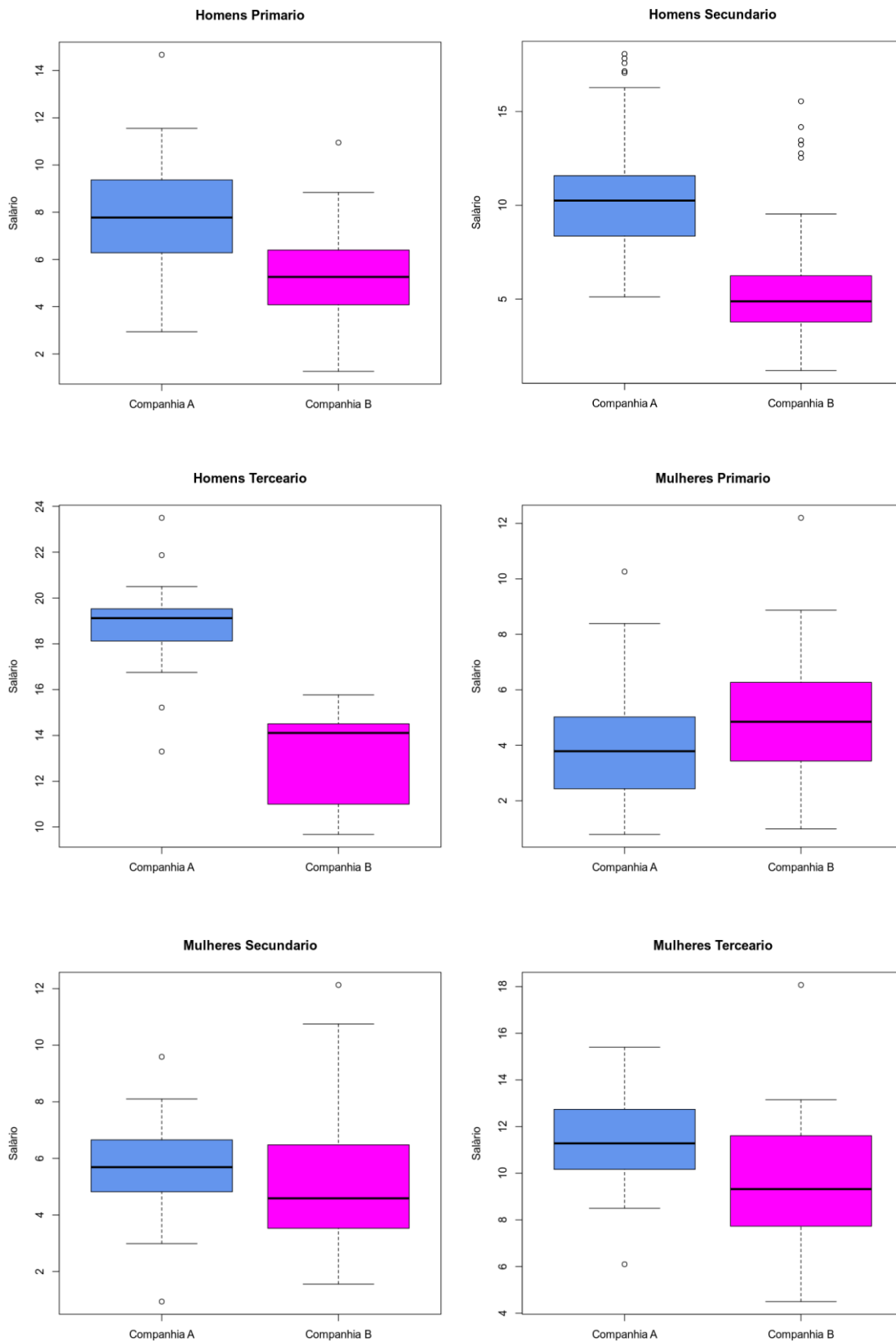
Tanto para a companhia A quanto para companhia B, através da análise do Box Plot de Distribuição Conjunta de Cargo Ocupado segundo Renda na companhia A e Box Plot de Distribuição Conjunta de Cargo Ocupado segundo Renda na companhia B o grupo constatou a existência de dependência entre as variáveis. As quais utilizando a média das variâncias, obteve-se para companhia A $\frac{var(S)}{var(s)} = \frac{\sum_{i=1}^k n_i var_i(s)}{\sum_{i=1}^k n_i} = 10,91$ e logo $R^2 = 1 - \frac{var(s)}{var(S)} = 0,53$ demonstrando uma associação de 53% entre as variáveis e, para companhia B, $\frac{var(S)}{var(s)} = 4,91$ e logo $R^2 = 0,48$ quantificando uma associação 48% para suas variáveis. Observa-se que em ambos os casos cargos gerenciais tendem a gerar uma renda substancialmente maior comparado a outros cargos, podendo essa ser mais que o dobro da primeira, porém observa-se também que grande maioria dos participantes possuem cargos não gerenciais, e, funcionários da companhia A tendem a ter maior remuneração que funcionários da companhia B.

vi) Conclusões sobre a análise de dependência

Analisando as variáveis citadas o grupo constatou que apenas os pares Escolaridade x Renda e Cargo Ocupado x Renda possuem dependência significativa para ambas as empresas, através da análise das variáveis infere-se também que a empresa A possivelmente trata-se de uma empresa com estilo mais tradicional e conservador devido a dependências entre gênero x cargo ocupado e gênero x salário observadas.

Parte 3

i) Comparação entre o salário para as categorias educação e gênero



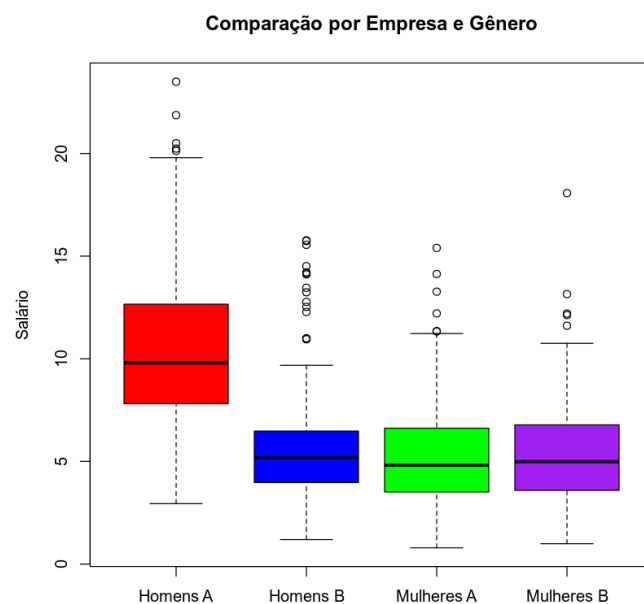
Através dos boxplots é possível observar que para todas as categorias, a média de salário é maior para a companhia A do que para a B, com exceção de mulheres cujo grau de

escolaridade é o ensino primário. Além disso, nesses casos o primeiro quartil é superior ao da empresa B, e o terceiro quartil é inferior. Quanto ao salário máximo, este é maior na empresa A para as categorias de homens e para mulheres com ensino terciário, e o menor salário para as categorias é sempre na companhia B, exceto para mulheres com o primário.

ii) Intervalo de Confiança para categorias de educação e gênero

alfa = 0.95	Homem		Mulher	
Primário	[1.775032 ; 3.31371]	0.7693393	[-1.481575 ; 0.02827959]	0.7549275
Secundário	[3.898977 ; 5.899783]	1.000403	[-0.3385604 ; 1.348969]	0.8437645
Terciário	[3.813117 ; 7.566661]	1.876772	[-1.82969 ; 4.380023]	3.104856

iv) Diferença salarial



v) Diferença Salarial hipótese

Hipóteses:

- $H_0: \mu_1 - \mu_2 = 0$
- $H_a: \mu_1 - \mu_2 \neq 0 \mid \alpha = \text{gama} = 0.99$

Média de salário Homem = Mulher Empresa A: Rejeição

Média de salário Homem = Mulher Empresa B: Aceitação

vi) Participação de mulheres em cargos gerenciais usando intervalos de confiança

Gerencia A = B

margem: 0.2309593

[-0.5367564 ; -0.07483781] Confiança: 0.95

vii) Participação de mulheres em cargos gerenciais usando teste de hipótese

Hipóteses:

- $H_0: \mu_1 - \mu_2 = 0$
- $H_a: \mu_1 - \mu_2 \neq 0 \mid \alpha = \text{gama} = 0.99$

Aceita H_0 com 0.99, médias IGUAIS

ret dentro do limite

ret: 2.707946 | limite 2.756386

Parte 4

i) Em ambas as companhias quanto maior o salário maior o grau de escolaridade, logo a maioria das pessoas que recebem os maiores salários tem grau de escolaridade “terciária”. Em contrapartida temos que quanto menor o salário, menor o grau de escolaridade. Na companhia A, de 129 pessoas com grau de escolaridade primário, apenas 20 recebem mais de 8,94 salários (15,5%), enquanto que das 42 com grau de escolaridade terciário, 40 recebem mais de 8,94 salários (95,2%). Na companhia B, podemos ver algo parecido, onde das 112 pessoas com primário, apenas 2 recebem mais de 9,85 salários (1,8%), enquanto que das 19 que tem terciário, 12 recebem mais de 9,85 salários (63,2%).

ii) Na companhia A, existem 12 mulheres e 34 homens na gerência. Na companhia B, em contrapartida, temos 17 mulheres e 13 homens na gerência, ou seja, existem mais mulheres na gerência na Companhia B (57%) do que na Companhia A (26%) levando em consideração o número de cargos gerenciais em cada companhia.

iii) Na companhia A, das 126 mulheres, 4 recebem mais de 11,90 salários, o que é 3% do total de mulheres e 1,3% do total de pessoas, enquanto que dos 173 homens, 48 recebem mais de 11,90 salários, o que é 28% dos homens e 16,1% do total de pessoas. Na companhia B, das 118 mulheres, 7 recebem mais de 9,85 salários, o que é 6% do total de mulheres e 2,8% do total de pessoas, enquanto que dos 129 homens, 15 recebem mais de 9,85 salários, o que é 12% do total de homens e 6% do total de pessoas. Assim, podemos concluir que, as mulheres recebem menos que os homens em ambas as empresas e isso é mais evidente na companhia A do que na companhia B, onde o valor é mais distribuído.

v) Na companhia A o maior salário é 23,50, a média é 8,58, a mediana é 7,78 e a moda é 4,58. 81 das 299 (27% do total) das pessoas recebem salário no intervalo 5,99 – 8,94 e 18% das pessoas recebem mais que a metade do maior salário da companhia (55 / 299). Na companhia B, por sua vez, o maior salário é 18,07, a média é 5,66, a mediana é 5,10 e a moda é 5,06. 85 das 247 pessoas (34% do total), recebem salário no intervalo 3,98 – 5,94 e 12% das pessoas recebem mais da metade do maior salário da companhia (30 / 247). Portanto, a companhia A paga melhores salários.