Universidade Estadual de Campinas
Instituto de Matemática, Estatística e Computação Cientifica
Departamento de Estatística

Relatório - Parte I
Exercício 1

Guilherme Pazian RA:160323
Henrique Capatto RA:146406
Hugo Calegari RA:155738
Leonardo Uchoa Pedreira RA:156231

Professor: Caio Lucidius Naberezny Azevedo

Campinas-SP, 12 de Junho de 2017

Exercício 1

SUGESTÃO:
O código do trabalho poderia terminar antes do início do trabalho de forma a melhorar a vizualização e organização

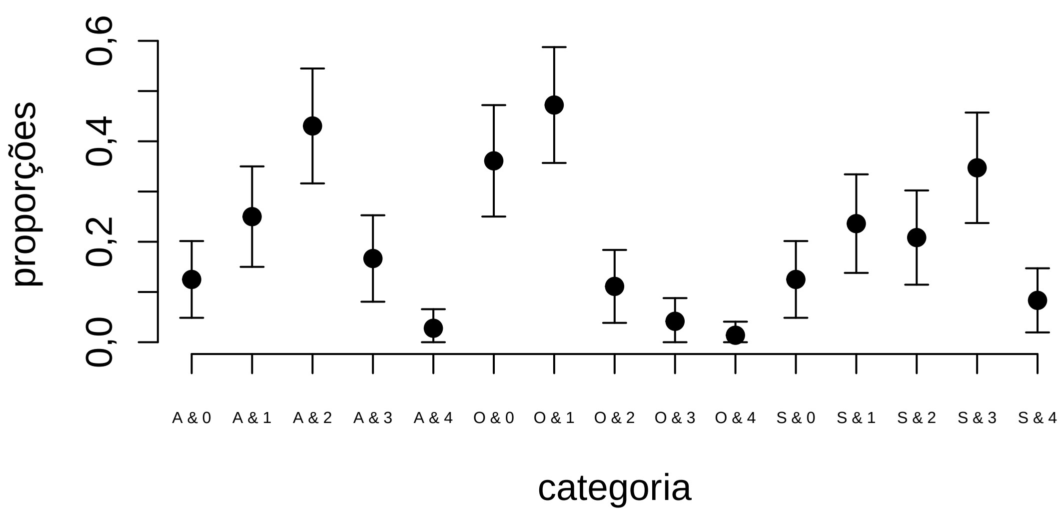
1.Introdução

Descrição do problema, conjunto de dados e objetivos.

2. Análise Descritiva

Toda a parte da análise descritiva, incluindo gráficos pertinentes.

	Nível de microinfiltração					
Material	0	1	2	3	4	Total
Allbond	9,00	18,00	31,00	12,00	2,00	72,00
Optibond	26,00	34,00	8,00	3,00	1,00	72,00
Scotchbond	9,00	17,00	15,00	25,00	6,00	72,00



3. Análise Inferencial

Objetivos: Descrição do(s) modelo(s), análise(s) de resíduo(s), comparações de interesse, gráficos e comentários (a escolha dos níveis de significância fica à cargo de cada equipe, devendo os valores adotadas, serem informados no relatório). Naturalmente, quando determinado, deverão ser usadas as metodologias constantes na questão. Caso a metodologia (modelo) usado não se adeque bem aos dados, comentários a respeito deverão ser feitos, mencionando que outras metodologias devem ser utilizadas (não, necessariamente, precisa ser dito qual(is)).

O Modelo probabílistico gerador da Tabela é o Produto de Multinomiais Independentes pois os totais marginais relacionados aos tipos de materiais de Selante são fixados, portanto o modelo é dado pela seguinte equação:

$$N_i = (N_{i0}, N_{i1}, N_{i2}, N_{i3}, N_{i4})' \sim multinomial(n_{i.}, \theta_i) \;\; i = 1,2,3$$

$N_i \perp N_j \; \forall \; i \neq j$; $i, j = 1,2,3$, $\theta_{\mathbf{i}} = (\theta_{i0}, \theta_{i1}, \theta_{i2}, \theta_{i3})'$ e $\theta_{ij} \in (0,1)$

Matriz

Medida de Desempenho:

A medida de desempenho proposta foi a média dos escores ponderadas pela probabilidade de classificação em cada grupo. Temos então que as medidas de desempenho dos materiais são dadas por:

$\bar{S}_{Allbond} = 0\theta_{10} + 1\theta_{11} + 2\theta_{12} + 3\theta_{13} + 4\theta_{14}$

$\bar{S}_{Optibond} = 0\theta_{20} + 1\theta_{21} + 2\theta_{22} + 3\theta_{23} + 4\theta_{24}$

$\bar{S}_{Scotchbond} = 0\theta_{30} + 1\theta_{31} + 2\theta_{32} + 3\theta_{33} + 4\theta_{34}$

Temos o interesse em testar se as medidas de desempenho dos três materiais são iguais, ou seja, testar a hipóteses:

$$H_0 : \begin{cases} \bar{S}_{Allbond} - \bar{S}_{Optibond} = 0 \\ \bar{S}_{Allbond} - \bar{S}_{Scotchbond} = 0 \end{cases} \; Vs \; H_1 : \text{Há pelo menos uma diferença}$$

A qual é equivalente a testar:

$$H_0 : \; B\pi = D \;\; Vs \;\; H_1 : B\pi \neq D$$

onde:

$$B^{'} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 2 & 2 \\ 3 & 3 \\ 4 & 4 \\ 0 & 0 \\ -1 & 0 \\ -2 & 0 \\ -3 & 0 \\ -4 & 0 \\ 0 & 0 \\ 0 & -1 \\ 0 & -2 \\ 0 & -3 \\ 0 & -4 \end{bmatrix} ; \pi = \begin{bmatrix} \theta_{10} \\ \theta_{11} \\ \theta_{12} \\ \theta_{13} \\ \theta_{14} \\ \theta_{20} \\ \theta_{21} \\ \theta_{22} \\ \theta_{23} \\ \theta_{24} \\ \theta_{30} \\ \theta_{31} \\ \theta_{32} \\ \theta_{33} \\ \theta_{34} \end{bmatrix} \; e \;\; D = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Conforme Azevedo(XXXano), podemos testar as hipóteses acima utilizando um teste assintótico baseado na estatística qui-quadrado. Ao realizar o teste, observou-se um valor de 54,27 que , respeitando-se os graus de liberdade da estatística do teste ("2" nesse caso), obtemos um p-valor de <0.0001, ou seja, temos evidências contra a hipóte nula. Desta maneira, temos evidências contra a hipótese de que as medidas de desempenho dos três materiais são iguais. Uma vez que temos a indicação de que as medidas de desempenho dos materiais são diferentes, podemos agora testar as hipóteses de igualdade dos materiais dois a dois. Temos então as novas hipóteses:

$$H_{01} : \bar{S}_{Allbond} - \bar{S}_{Optibond} = 0 \; Vs \; H_{11} : \bar{S}_{Allbond} - \bar{S}_{Optibond} \neq 0$$

A qual é equivalente a testar:

$$H_{01} : \; B_1\pi = D \;\; Vs \;\; H_{11} : B_1\pi \neq D$$

Os vetores B_1, π e D estão definidos em anexo.

A partir do teste da hipótese acima, observou-se um valor de 30,42 que , respeitando-se os graus de liberdade da estatística do teste ("1" nesse caso), obtemos um p-valor de <0.0001

$$H_{02} : \bar{S}_{Allbond} - \bar{S}_{Scotchbond} = 0 = 0 \; Vs \; H_{12} : \bar{S}_{Allbond} - \bar{S}_{Scotchbond} \neq 0$$

A qual é equivalente a testar:

$$H_{02}: \ B_2\pi = D \ \ Vs \ \ H_{12}: B_2\pi \neq D$$

Os vetores B_2, π e D estão definidos em anexo. A partir do teste da hipótese acima, observou-se um valor de 2,84 que , respeitando-se os graus de liberdade da estatística do teste ("1" nesse caso), obtemos um p-valor de 0,0919

$$H_{03}: \bar{S}_{Optibond} - \bar{S}_{Scotchbond} = 0 = 0 \ Vs \ H_{13}: \bar{S}_{Optibond} - \bar{S}_{Scotchbond} \neq 0$$

A qual é equivalente a testar:

$$H_{03}: \ B_3\pi = D \ \ Vs \ \ H_{13}: B_3\pi \neq D$$

Os vetores B_3, π e D estão definidos em anexo.

Os vetores B_2, π e D estão definidos em anexo. A partir do teste da hipótese acima, observou-se um valor de 44,21 que , respeitando-se os graus de liberdade da estatística do teste ("1" nesse caso), obtemos um p-valor de <0.0001.

Portanto, temos a indicação de que as medidas de desempenho dos três materiais diferem estatísticamente entre si.

Uma outra análise poderia ser feita a partir de um modelo de regressão linear para Tabela de Contingência, escrito na forma $A\pi = X\beta$.

Dados os resultados expostos anteriormente, podemos testar o ajuste de um modelo que leva em consideração a hipótese de que todos as medidas de desempenho dos materiais são diferentes, e partir desse modelo se obtem as estimativas para a medida de desempenho de cada material.

Pode-se observar, através da tabela XXX os valores estimados da medida de desempenho de cada medida, seus respectivos desvios padrão (DP) e respectivos intervalos de confiança podem ser visualizados na tabela abaixo

Para fácil identificação, os nomes das colunas indicando as quantidades calculadas foi renomeada para melhor visualização. DP significa desvio Padrão, LIIC significa Limite Inferior do Intervalo de Confiança(IC) e LSIC, Limite Superior do IC.

Material	MD Estimada	DP	LIIC	LSIC
Allbond	1.72	0.11	1.50	1.95
Optibond	0.87	0.10	0.68	1.07
Scotchbond	2.03	0.14	1.75	2.30

A estatística **acho que não precisa esse modelo, reservar caso falte algo** Depois, utilizando a medida de desempenho para Allbond como referência, compararemos as outras duas medidas em relação aquela para determinação de qual é o melhor material

Já para o segundo modelo, temos a seguinte tabela

	Estimado	LIIC	LSIC
1	0.85	0.55	1.15
2	-0.31	-0.66	0.05

4. Conclusões

O que se pode concluir da análise, em termos do problema apresentado, e críticas em relação a análise feita.

5. Anexos

Para testar a hipótese

$$H_{01}: \bar{S}_{Allbond} - \bar{S}_{Optibond} = 0 \; Vs \; H_{11}: \bar{S}_{Allbond} - \bar{S}_{Optibond} \neq 0 \leftrightarrow H_{01}: \; B\pi = D \;\; Vs \;\; H_{11}: B\pi \neq D$$

definimos:

$$B_1^{'} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 0 \\ -1 \\ -2 \\ -3 \\ -4 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} ; \pi = \begin{bmatrix} \theta_{10} \\ \theta_{11} \\ \theta_{12} \\ \theta_{13} \\ \theta_{14} \\ \theta_{20} \\ \theta_{21} \\ \theta_{22} \\ \theta_{23} \\ \theta_{24} \\ \theta_{30} \\ \theta_{31} \\ \theta_{32} \\ \theta_{33} \\ \theta_{34} \end{bmatrix} \; e \;\; D = \begin{bmatrix} 0 \end{bmatrix}.$$

Para testar a hipótese:

$$H_{02}: \bar{S}_{Allbond} - \bar{S}_{Scotchbond} = 0 = 0 \; Vs \; H_{12}: \bar{S}_{Allbond} - \bar{S}_{Scotchbond} \neq 0 \leftrightarrow H_{02}: \; B\pi = D \;\; Vs \;\; H_{12}: B\pi \neq D$$

$$B_2^{'} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ -2 \\ -3 \\ -4 \end{bmatrix} ; \pi = \begin{bmatrix} \theta_{10} \\ \theta_{11} \\ \theta_{12} \\ \theta_{13} \\ \theta_{14} \\ \theta_{20} \\ \theta_{21} \\ \theta_{22} \\ \theta_{23} \\ \theta_{24} \\ \theta_{30} \\ \theta_{31} \\ \theta_{32} \\ \theta_{33} \\ \theta_{34} \end{bmatrix} \; e \;\; D = \begin{bmatrix} 0 \end{bmatrix}.$$

Para testar a hipótese:

$$H_{03} : \bar{S}_{Optibond} - \bar{S}_{Scotchbond} = 0 = 0 \; Vs \; H_{13} : \bar{S}_{Optibond} - \bar{S}_{Scotchbond} \neq 0 \leftrightarrow H_{03} : \; B\pi = D \;\; Vs \;\; H_{13} : B\pi \neq D$$

onde:

$$B_3^{'} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 0 \\ -1 \\ -2 \\ -3 \\ -4 \end{bmatrix} ; \pi = \begin{bmatrix} \theta_{10} \\ \theta_{11} \\ \theta_{12} \\ \theta_{13} \\ \theta_{14} \\ \theta_{20} \\ \theta_{21} \\ \theta_{22} \\ \theta_{23} \\ \theta_{24} \\ \theta_{30} \\ \theta_{31} \\ \theta_{32} \\ \theta_{33} \\ \theta_{34} \end{bmatrix} \; e \;\; D = \begin{bmatrix} 0 \end{bmatrix}.$$

Para os modelos de regressão temos as seguintes hipóteses:

$$H_0 : \begin{cases} \bar{S}_{Allbond} = \alpha_1 \\ \bar{S}_{Optibond} = \alpha_2 \\ \bar{S}_{Scotchbond} = \alpha_3 \end{cases} \; Vs \; H_1 : \text{Há pelo menos uma diferença}$$

Temos então o modelo de regressão definido como