

Natalia Elis Giordani, Prof. Dr. João Riboldi

Departamento de Estatística - UFRGS

## Introdução

A análise de variância (Anova) é geralmente robusta em relação a violações das suas suposições de aditividade, homogeneidade de variâncias e normalidade. No entanto, a Anova perde robustez quando a suposição de independência é violada, podendo trazer fortes consequências por introduzir um viés que se mostra muitas vezes significativo.

## Objetivo

Busca-se neste trabalho, verificar as consequências da violação da suposição de independência na análise de variância sob o ponto de vista de consequências no valor de probabilidade e poder no teste F da Anova, considerando diferentes formas de não-independência, utilizando-se simulação de dados.

## Metodologia

Considerou-se um modelo de Anova em classificação simples, com diferentes estruturas de correlação, efeitos nulos e não-nulos de tratamentos, distribuição Normal para o erro e correlação devido a grupos e a seqüência. Utilizando o software SAS simulou-se 10000 experimentos para cada caso.

Para correlação devido a grupos considerou-se quatro tratamentos com seis observações por tratamento, correlação entre todas as observações e entre observações de cada grupo de forma igual ou diferente. Os grupos foram considerados na simulação, na análise, ou em ambos. Para correlação devido a seqüência, considerou-se três tratamentos com números diferentes de observações por tratamento, procedendo-se medidas repetidas sobre a mesma unidade experimental em sete tempos diferentes. Foram consideradas oito estruturas de covariância para os dados (Independente, CS, UN, DVD, AR, H-F, Toep e Toep Banded).

Para cada experimento simulado realizou-se o teste F da Anova e calculou-se os níveis mínimos de significância (nms). Para o caso de efeitos nulos de tratamentos avaliou-se a aderência a distribuição Uniforme (0,1), utilizando-se os testes de Kolmogorov-Smirnov, Cramér-von Mises e Anderson-Darling. No caso de efeitos não-nulos, avaliou-se o poder através da distribuição de freqüência dos nms.

## Resultados

### Correlação devido a grupos

TABELA 1 – Precisão e Poder considerando Correlação Devido a Grupos

Estrutura	Correlação	Precisão				Poder	
		K-S	C-M	A-D	AS	S	NS
Correlação entre todas observações, grupos não considerados		<0,001	<0,001	<0,001	8937	454	609
Correlação entre observações dentro de cada grupo, grupos considerados na simulação		<0,001	<0,001	<0,001	4947	2887	2166
Correlação entre observações dentro de cada grupo, grupos considerados na simulação e na análise	Mesmo valor	<0,001	<0,001	<0,001	4895	2223	2882
Correlação entre todas observações, grupos considerados na simulação		<0,001	<0,001	<0,001	8937	454	609
Correlação entre todas observações, grupos considerados na simulação e na análise		<0,001	<0,001	<0,001	8412	777	811
Correlação entre todas observações, grupos não considerados		<0,001	<0,001	<0,001	9964	23	13
Correlação entre observações dentro de cada grupo, grupos considerados na simulação		<0,001	<0,001	<0,001	2138	3908	3954
Correlação entre observações dentro de cada grupo, grupos considerados na simulação e na análise	Valores diferentes	>0,250	>0,250	>0,250	2138	3908	3954
Correlação entre todas observações, grupos considerados na simulação		<0,001	<0,001	<0,001	9396	423	181
Correlação entre todas observações, grupos considerados na simulação e na análise		<0,001	<0,001	<0,001	3723	3783	2494

### Correlação devido à seqüência

TABELA 2 – Precisão e Poder considerando Correlação devido a Seqüência

Estrutura de Covariância	Efeito	Precisão				Poder	
		K-S	C-M	A-D	AS	S	NS
Independente	Tempo	>0,250	0,192	0,214	1379	1874	6747
	Tratamento	0,01	0,03	0,042	9997	3	0
	Tempo*Trat	>0,250	>0,250	>0,250	678	1238	8084
Componente Simétrico	Tempo	<0,001	<0,01	<0,01	9226	375	399
	Tratamento	<0,001	<0,01	<0,01	9653	140	207
	Tempo*Trat	<0,001	<0,01	<0,01	9925	48	27
Não-Estruturada	Tempo	<0,001	<0,01	<0,01	9240	350	410
	Tratamento	<0,001	<0,01	<0,01	9638	160	202
	Tempo*Trat	<0,001	<0,01	<0,01	9929	52	19
Variâncias Heterogêneas	Tempo	0,188	0,147	0,165	1396	1846	6758
	Tratamento	0,029	0,034	0,044	9997	3	0
	Tempo*Trat	>0,250	>0,250	>0,250	677	1250	8073
Auto-Regressiva	Tempo	<0,001	<0,01	<0,01	9209	379	412
	Tratamento	<0,001	<0,01	<0,01	9653	152	195
	Tempo*Trat	<0,001	<0,01	<0,01	9926	50	24
H-F	Tempo	<0,001	<0,01	<0,01	9221	370	409
	Tratamento	<0,001	<0,01	<0,01	9434	239	327
	Tempo*Trat	<0,001	<0,01	<0,01	9926	47	27
Toeplitz	Tempo	<0,001	<0,01	<0,01	9230	372	398
	Tratamento	<0,001	<0,01	<0,01	9659	147	194
	Tempo*Trat	<0,001	<0,01	<0,01	9933	46	21
Toeplitz Bandeada	Tempo	<0,001	<0,01	<0,01	5109	1869	3022
	Tratamento	<0,001	<0,01	<0,01	9997	3	0
	Tempo*Trat	<0,001	<0,01	<0,01	5271	1988	2741

## Conclusão

**Considerando correlação devido a grupos, o teste F se mostrou impreciso. Quanto ao poder percebe-se dependência em função da estrutura de correlação.**

**Para correlação devido à seqüência, quanto à precisão, percebe-se uma robustez em relação a variâncias diferentes, mas não a covariâncias diferentes. Quanto ao poder, percebe-se que as estruturas que contemplam covariância parecem não ser afetadas.**