

Análise de Covariância (ANCOVA)

Paulo S. P. Silveira (silveira@usp.br)

José O. Siqueira (siqueira@usp.br)

Contents

Objetivos	1
ANCOVA independente	1
Exemplo	3
Estatística descritiva	3
Regressões lineares simples	12
Suposições	16
Teste do efeito do fator	18
Tamanhos de efeito	20
Médias marginais estimadas ajustadas	20
Testes <i>post hoc</i>	25
ANCOVA mista	27
Exemplo	27
Planejamento	41
Referências	43

v20211212.1746

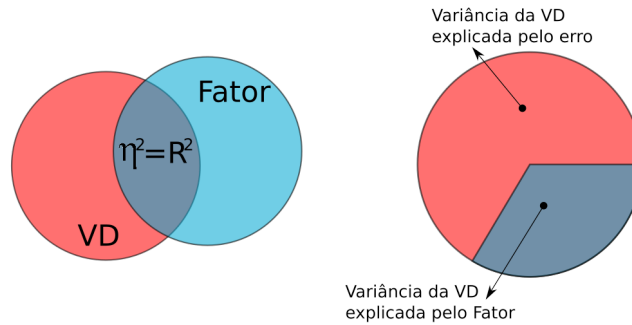
Objetivos

- Conceituar ANCOVA, distinguindo-a da ANOVA e da Regressão Linear Simples.
- Discorrer sobre as suposições e condições nas quais ANCOVA pode ser utilizada.
- Descrever e indicar dois delineamentos diferentes da ANCOVA:
 - com um fator entre participantes e uma covariável intervalar,
 - em delineamento pré-pós intervenção.
- Proceder com a análise descritiva numérica e gráfica dos dados.
- Formular e implementar os modelos de ANCOVA adequados a cada situação.
- Determinar as significâncias estatística e prática das suposições e do efeito do fator.

ANCOVA independente

ANCOVA é uma combinação de regressão linear simples com ANOVA unifatorial independente. Este é o primeiro passo para o modelo linear geral (GLM, *general linear model*).

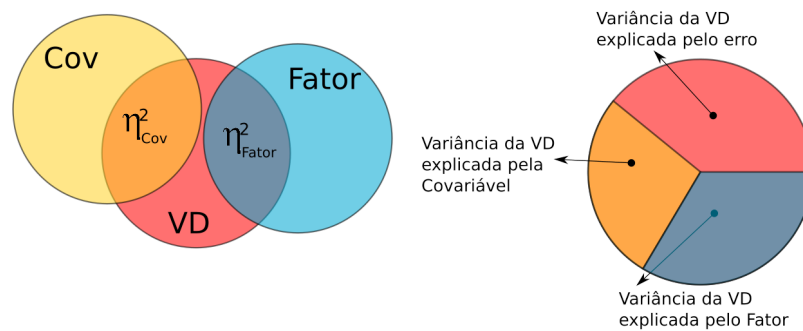
Na ANOVA o tamanho de efeito global, R^2 , é a porcentagem da variância da variável de desfecho (VD), que é intervalar, explicada pelo fator, que é nominal. Note que, como só existe um fator, o tamanho de efeito global coincide com o tamanho de efeito do fator, η^2 . A parcela da variância da VD que não é explicada pelo modelo da ANOVA (chamado de resíduo ou termo de erro) corresponde a $1 - \eta^2$.



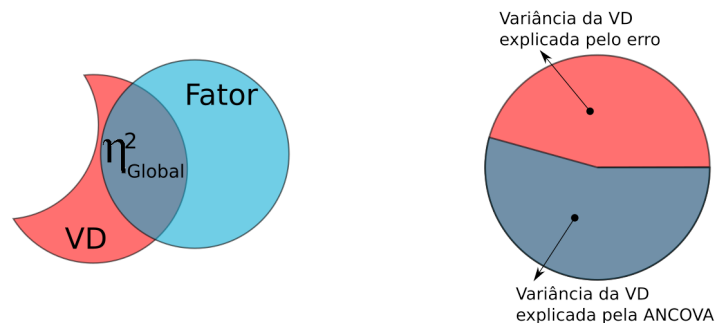
Existem duas razões principais para usar ANCOVA:

1. Reduzir a variância do resíduo:

A ANCOVA introduz uma covariável intervalar que também explica parcialmente a variância da VD. Supõe-se que esta covariável não esteja correlacionada com o fator em estudo. Com isso, a variância da VD é particionada em dois tamanhos de efeito parciais, uma parte atribuída ao fator, η^2_{Fator} e outra parte associada com a covariável, η^2_{Cov} . São estes tamanhos de efeito parciais que aparecerão na análise, adiante.



A introdução desta covariável permite o controlar melhor o efeito do fator. É como se a parcela da variância da VD atribuída à covariável fosse subtraída pelo modelo.



Como os tamanhos de efeitos são porcentagens da variância da VD explicadas pela covariável ou pelo fator, subtraindo-se a parte atribuída à covariável, altera-se o denominador. Com isto, a porcentagem remanescente explicada pelo fator é proporcionalmente maior do que aquela computada pela ANOVA. O modelo ganha em poder explicativo, e este é o tamanho de efeito global, denotado por η^2_{Global} ou simplesmente η^2 que aparece nas análises adiante.

Pela ilustração também é fácil de perceber que a soma dos tamanhos de efeito parciais não correspondem necessariamente ao tamanho de efeito global.

2. Ajustar as médias da covariável, de modo que o valor médio da covariável seja o mesmo em todos os grupos e depois ajustar as médias e erros-padrão da VD nos grupos (equalização

estatística de condições entre participantes):

O uso de uma covariável é uma forma de obter controle estatístico quando o controle experimental, obtido pelo delineamento do estudo, não for possível. Testes mais simples, como teste t ou ANOVA, muitas vezes podem ser usados em situações laboratoriais. Em um experimento com camundongos, havendo recursos financeiros, o pesquisador pode ter, por exemplo, animais com a mesma idade ou até mesmo com a mesma genética.

Suponha, no entanto, que os animais disponíveis têm idades variadas. Controlar o efeito do experimento pela idade pode ser uma boa ideia. Críticos da ANCOVA dirão que esta dificuldade poderia ser resolvida com randomização dos animais entre os grupos ou condições experimentais (na linguagem do GLM: Fator). Porém, randomização mitiga a discrepância entre as condições experimentais **em média** (com a intenção, neste exemplo, de equalizar as idades), mas não garante esta homogeneidade completamente em um determinado experimento, especialmente se a amostra não for grande.

Enquanto a ANCOVA pode ser preterida em relação à ANOVA em delineamentos experimentais nas quais o pesquisador assume que a randomização é suficiente, em estudos quase-experimentais ou observacionais, nos quais randomização não é realizada, ANCOVA é uma estratégia mais defensável.

O GLM é uma das formas mais simples para se operacionalizar este controle estatístico na análise para testar o efeito fixo do fator entre participantes.

Por “ajustar as médias da covariável, de modo que o valor médio da covariável seja o mesmo em todos os grupos”, o GLM computa a média da covariável em cada condição do fator, e então utiliza a média destas médias como a média representativa para a covariável.

O segundo passo é “ajustar as médias e erros-padrão da VD nos grupos (equalização estatística de condições entre participantes)”, que é alterar os valores das médias da VD para cada uma das condições do fator, o que é conhecido como *médias marginais ajustadas*, que também aparecerão nas análises da ANCOVA. Da mesma forma, os erros-padrão são homogeneizados. Veremos, adiante, como estes valores de média e erro-padrão são computados após a aplicação do modelo.

Exemplo

Um experimento foi projetado para descobrir se álcool (nas condições placebo, ingestão em pequena quantidade e ingestão em grande quantidade) afeta o desempenho de dirigir, medido por erros cometidos no simulador veicular. A expectativa é que o número de erros aumente quanto maior for a intoxicação pelo álcool.

Dancey & Reidy (2019, página 424)

No entanto, uma suposição perfeitamente razoável é que a experiência de direção também está relacionada aos erros cometidos ao dirigir, mesmo num simulador. Quanto mais experiente for o motorista, menor será a quantidade de erros. Assim, a experiência do motorista está negativamente associada aos erros ao dirigir.

O delineamento é entre participantes. O pesquisador randomizou voluntários entre os três grupos mas, ainda assim, gostaria de mitigar o efeito da experiência dos motoristas para estudar o efeito do álcool. O dados estão disponíveis em `AlcoholExperienceDrivingErrors.xlsx`.

Alguma estatística descritiva, os testes das suposições para a ANCOVA, a execução da ANCOVA e testes *post hoc* estão implementados em `demo_ANCOVA.R`.

Como é um *Rscript* longo, vamos executá-lo em parte.

Estatística descritiva

Implementado em `demo_ANCOVA_01descritiva.R`:

```
----  
Data  
----
```

ID	Grupo	ErrosDirecao	Experiencia
1	Placebo	5	12
2	Placebo	10	5
3	Placebo	7	9
4	Placebo	3	24
5	Placebo	5	15
6	Placebo	7	6
7	Placebo	11	3
8	Placebo	2	30
9	Placebo	3	20
10	Placebo	5	10
11	Placebo	6	7
12	Placebo	6	8
13	PoucoAlcool	5	21
14	PoucoAlcool	7	16
15	PoucoAlcool	9	7
16	PoucoAlcool	8	15
17	PoucoAlcool	2	30
18	PoucoAlcool	5	21
19	PoucoAlcool	6	12
20	PoucoAlcool	6	13
21	PoucoAlcool	4	26
22	PoucoAlcool	4	24
23	PoucoAlcool	8	9
24	PoucoAlcool	10	6
25	MuitoAlcool	8	29
26	MuitoAlcool	10	8
27	MuitoAlcool	8	26
28	MuitoAlcool	9	20
29	MuitoAlcool	11	18
30	MuitoAlcool	15	6
31	MuitoAlcool	7	12
32	MuitoAlcool	11	7
33	MuitoAlcool	8	15
34	MuitoAlcool	8	9
35	MuitoAlcool	17	3
36	MuitoAlcool	11	7

Descriptive Statistics

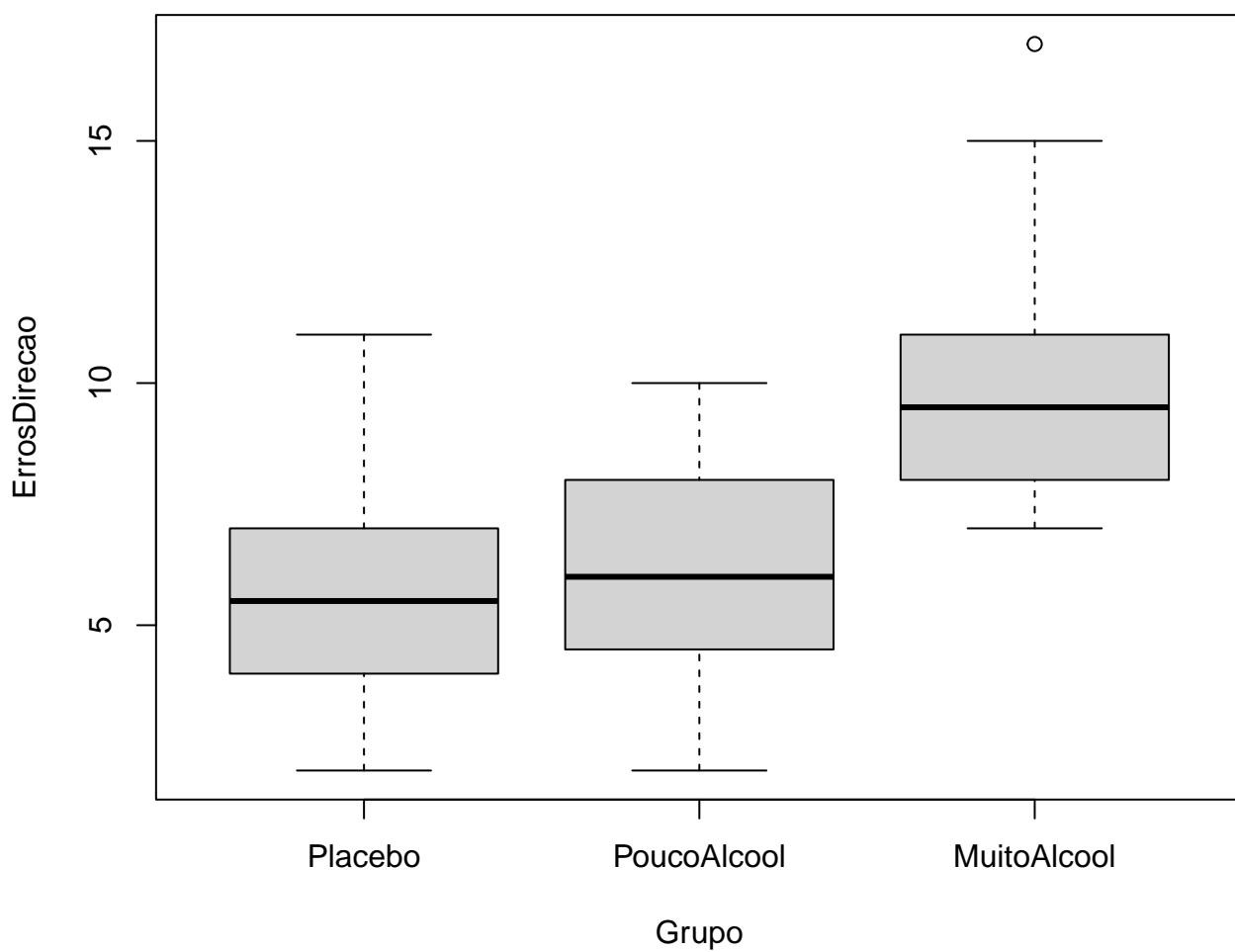
	Grupo		
ErrosDirecao	Placebo	PoucoAlcool	MuitoAlcool
2	1	1	0
3	2	0	0
4	0	2	0
5	3	2	0
6	2	2	0
7	2	1	1
8	0	2	4
9	0	1	1
10	1	1	1
11	1	0	3

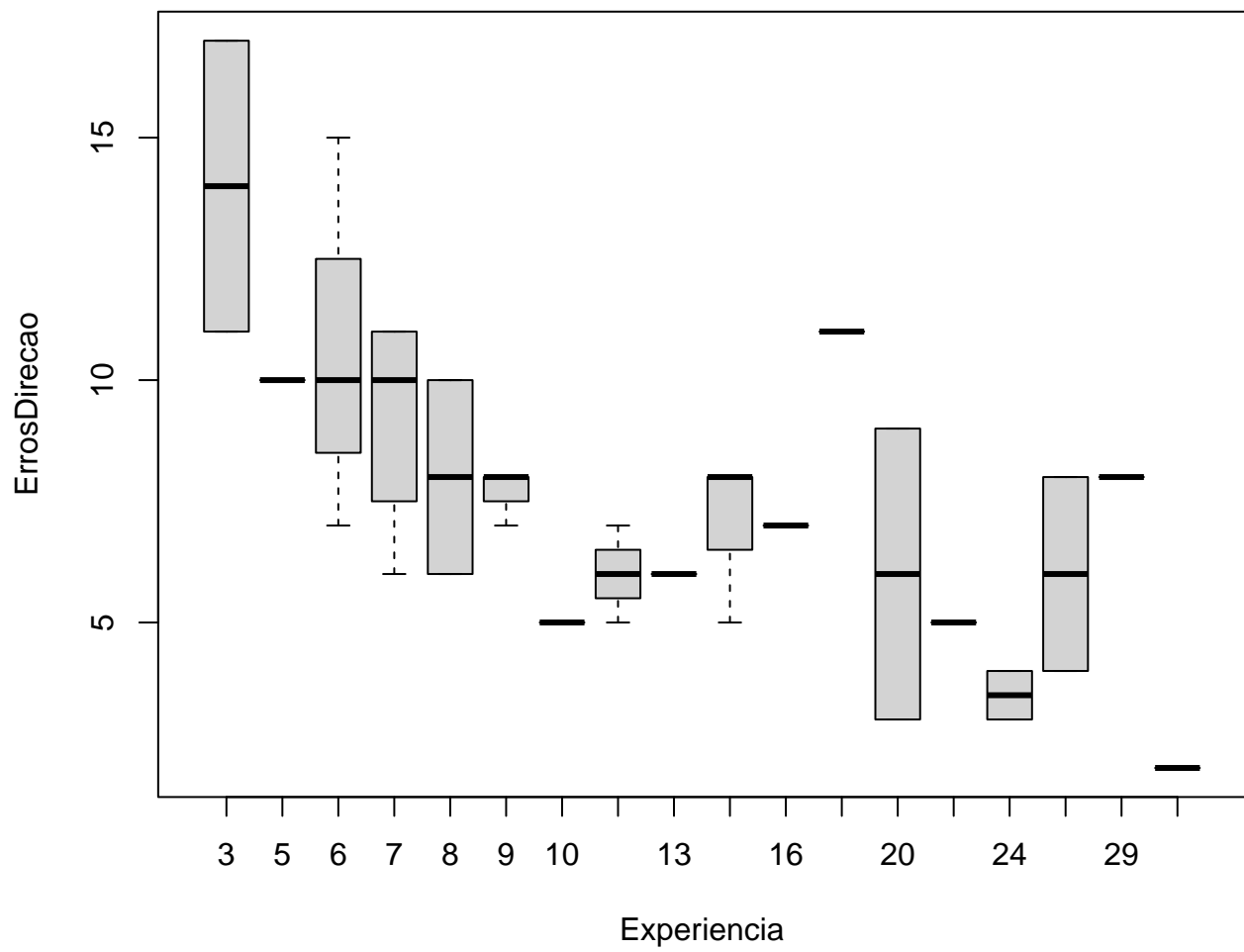
[illegible]

7	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	1
9	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
10	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0
15	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

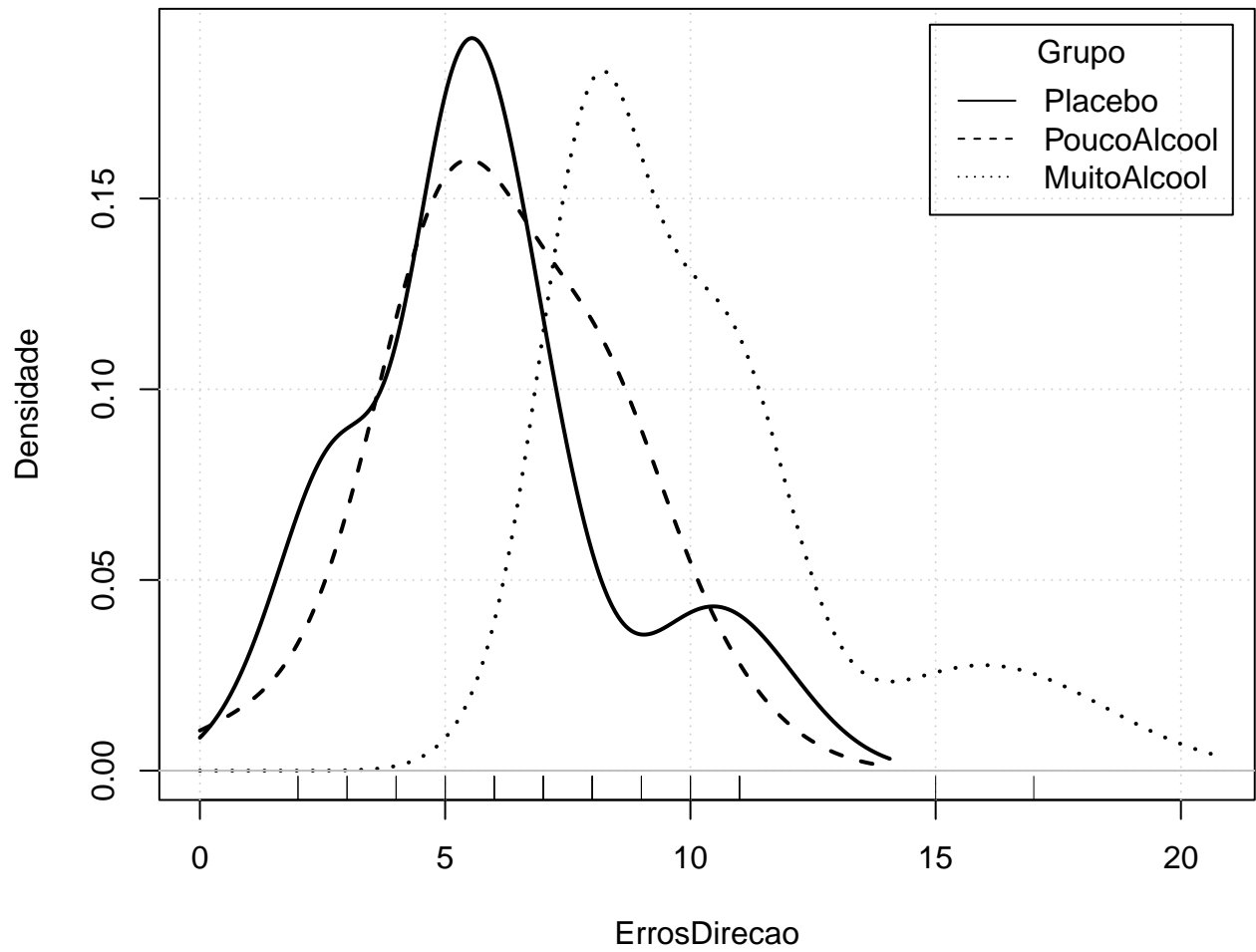
Sumario:

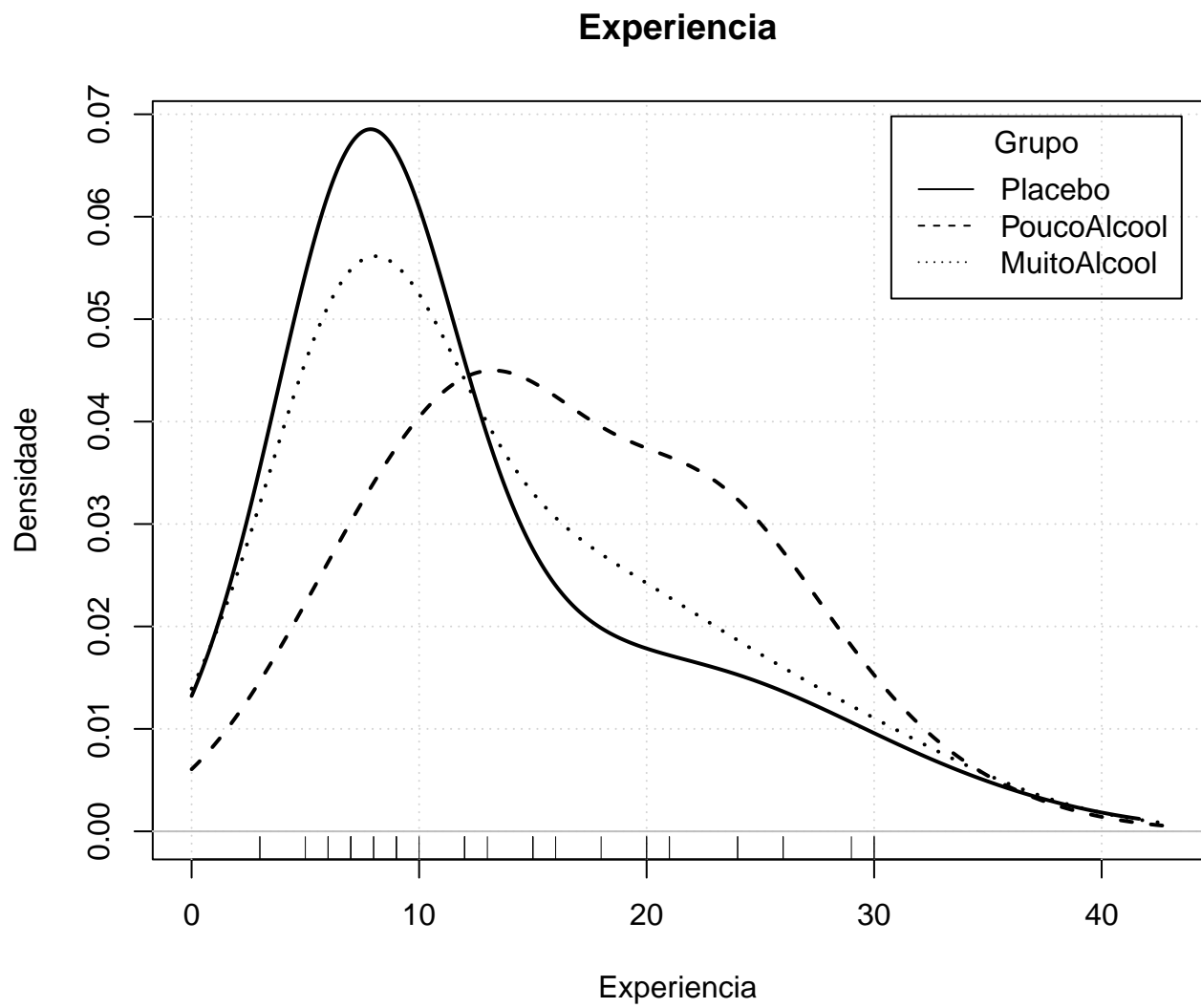
	Grupo	Media(ErrosDirecao)	SE(ErrosDirecao)
1	Placebo	5.833333	0.2242646
2	PoucoAlcool	6.166667	0.1940834
3	MuitoAlcool	10.250000	0.2540704

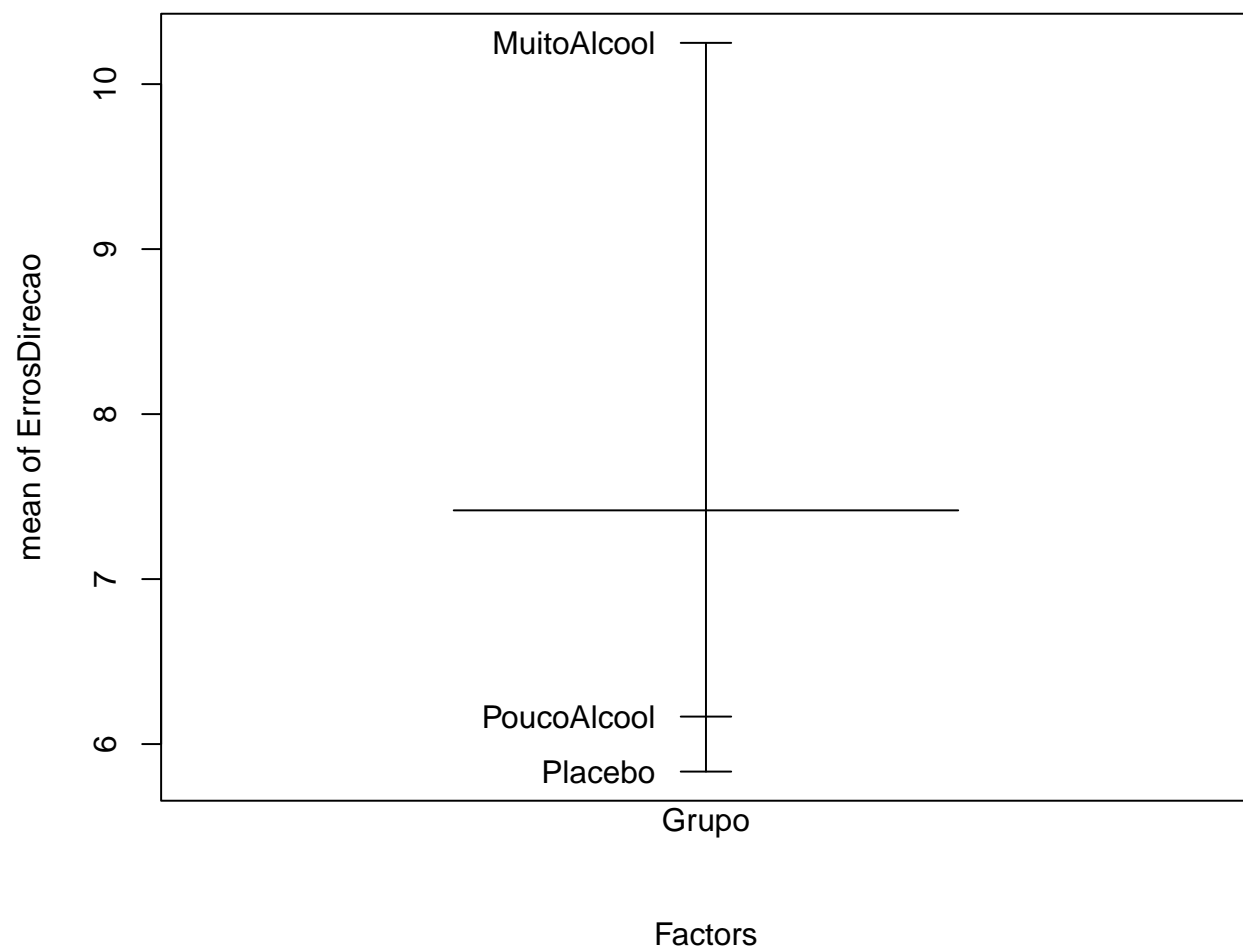


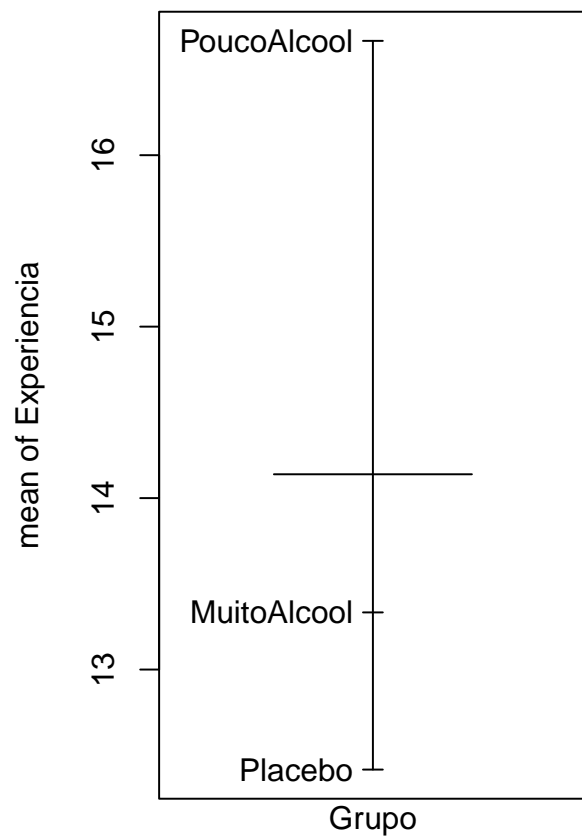


ErrosDirecao

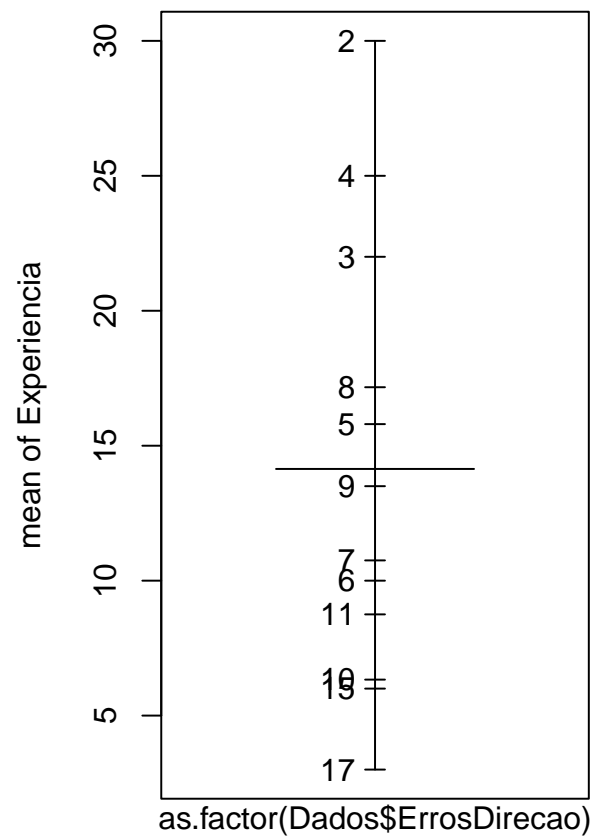




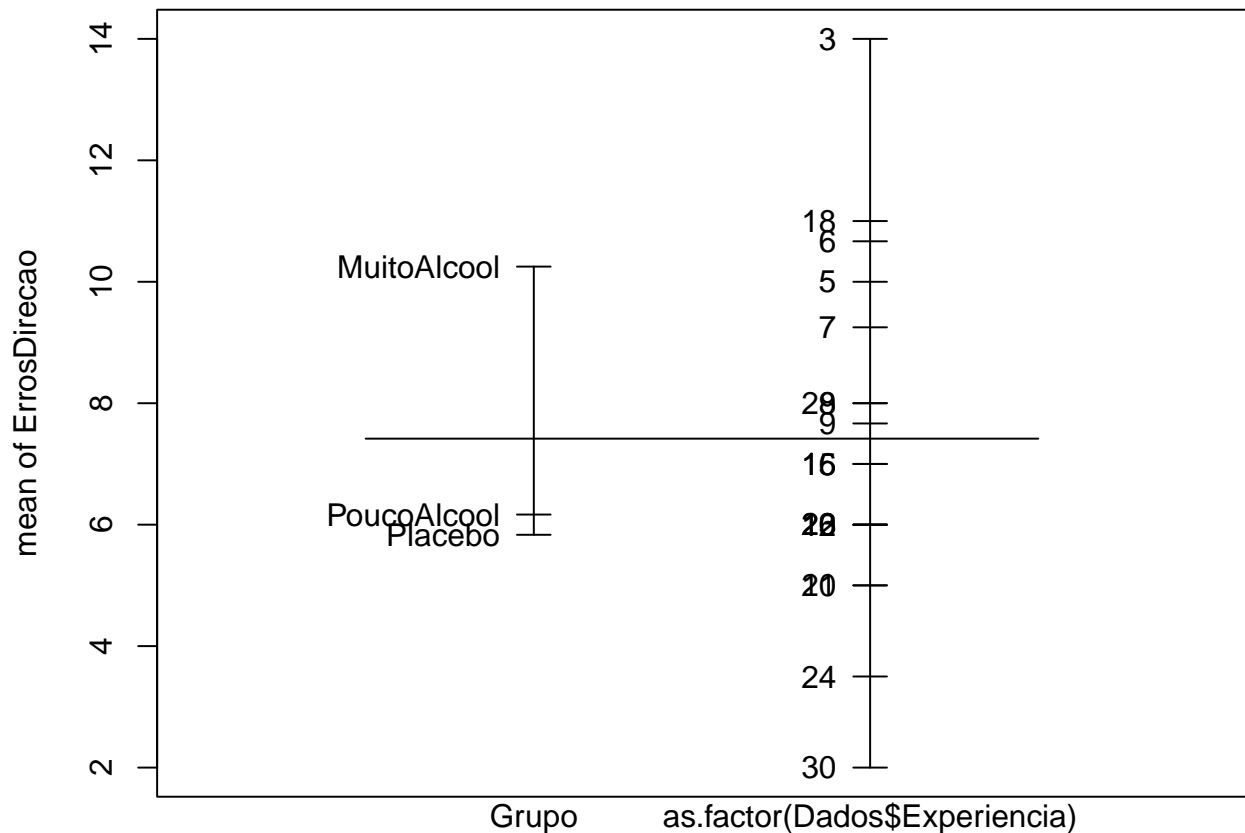




Factors



ErrosDirecao



Factors

O arquivo de dados contém as informações de cada voluntário (identificados pelo ID) em uma linha. O delineamento é completamente entre-participantes (indivíduos diferentes foram submetidos à ingestão de **Placebo**, **PoucoAlcool** e **MuitoAlcool**). Este placebo é uma bebida idealizada (neste exemplo fictício) com características idênticas em aspecto e sabor à bebida alcóolica, mas que não contém álcool.

Três tabelas mostram, respectivamente, em cada uma das condições do fator:

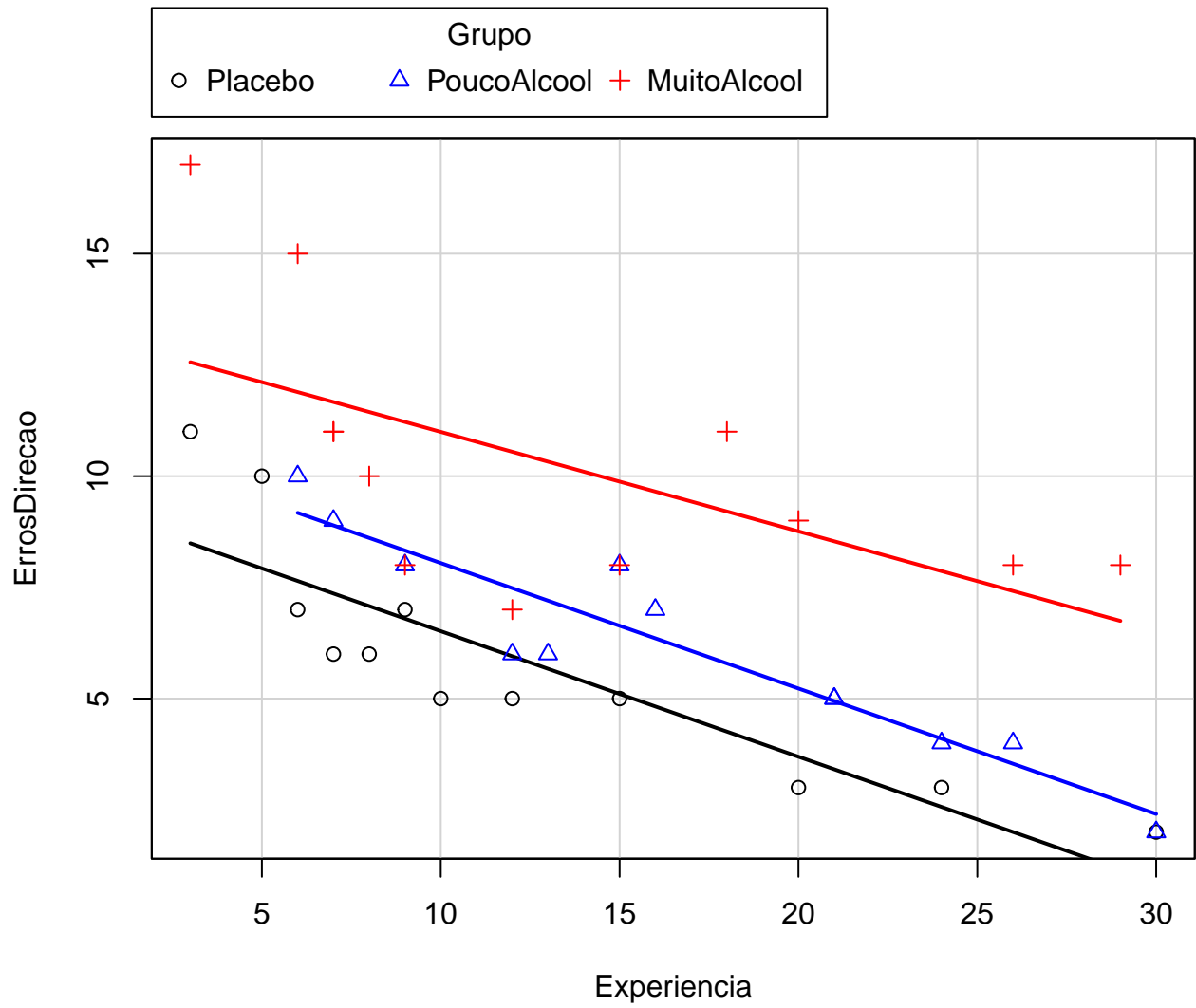
- a distribuição dos erros cometidos,
- a distribuição da experiência dos voluntários,
- a distribuição simultânea dos erro e experiência simultaneamente observados e
- um sumário com a média e erro padrão (**se**: *standard error*) dos erros cometidos.

Na sequência há vários gráficos para que possam escolher quais são adequados para o propósito da sua pesquisa.

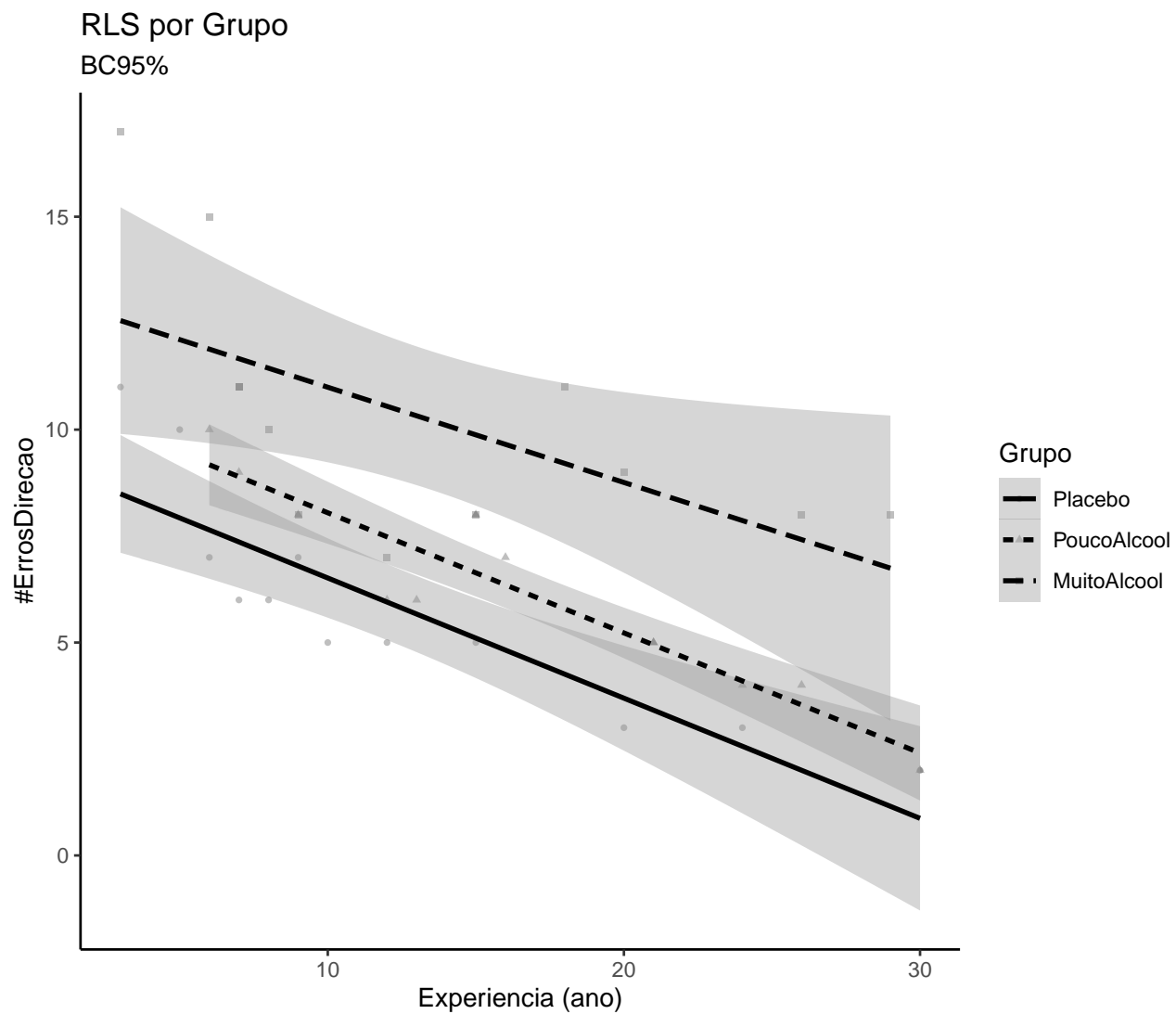
Regressões lineares simples

Implementado em `demo_ANCOVA_02lm.R`:

```
-----
Simple linear regressions
-----
```



`geom_smooth()` using formula 'y ~ x'



```
-----
Placebo
-----
```

Call:

```
lm(formula = ErrosDirecao ~ Experiencia, data = dt_tmp)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.5153	-0.9831	-0.3742	0.6088	2.5093

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	9.33732	0.74221	12.580	1.87e-07 ***
Experiencia	-0.28220	0.05034	-5.606	0.000226 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.387 on 10 degrees of freedom

```
Multiple R-squared:  0.7586,    Adjusted R-squared:  0.7345
F-statistic: 31.43 on 1 and 10 DF,  p-value: 0.0002257
```

```
-----
PoucoAlcool
-----
```

Call:

```
lm(formula = ErrosDirecao ~ Experiencia, data = dt_tmp)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.4829	-0.3483	0.0555	0.5106	1.3633

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	10.86731	0.59864	18.153	5.52e-09 ***
Experiencia	-0.28204	0.03281	-8.595	6.25e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8434 on 10 degrees of freedom

Multiple R-squared: 0.8808, Adjusted R-squared: 0.8689

F-statistic: 73.88 on 1 and 10 DF, p-value: 6.247e-06

```
-----
MuitoAlcool
-----
```

Call:

```
lm(formula = ErrosDirecao ~ Experiencia, data = dt_tmp)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.5482	-1.5513	-0.2127	1.3885	4.4392

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	13.23169	1.42182	9.306	3.06e-06 ***
Experiencia	-0.22363	0.09149	-2.444	0.0346 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.53 on 10 degrees of freedom

Multiple R-squared: 0.374, Adjusted R-squared: 0.3114

F-statistic: 5.974 on 1 and 10 DF, p-value: 0.0346

São apresentados dois tipos de gráfico. O primeiro é gerado pela função `car::scatterplot` que plota retas de regressão automaticamente, usando-se o parâmetro `regLine=TRUE`. O segundo usa recursos do pacote `ggplot2`, parametrizado para exibir as linhas de regressão com bandas de confiança 95%. Observe que as retas parecem paralelas, mas quando fazemos as bandas de confiança não conseguimos visualizar um segmento de reta comum às três retas, sugerindo que não há uma reta populacional de associação linear entre erros de

direção e experiência em direção que atenda aos três grupos (i.e., possivelmente o álcool distingue as três populações).

Temos, então, três regressões lineares simples, mostrando que há associação negativa significativa da experiência em direção com o número de erros cometidos no simulador em cada uma das condições do fator. Visualmente, as três regressões parecem paralelas, o que nos leva à discussão sobre as suposições necessárias para a aplicação da ANCOVA.

Suposições

Além das suposições da regressão linear simples e da ANOVA, a ANCOVA supõe:

- dissociação entre o fator e a covariável,
- covariável sem erro de mensuração (não testável, assumida pelo modelo),
- linearidade entre a VD e a covariável em todos os níveis do fator,
- homogeneidade da inclinação destas regressões lineares simples.

Implementado em `demo_ANCOVA_03suposicoes.R`:

```
-----
Assumptions
-----

-----
Dissociation between Grupo (factor) and Experiencia (covariate)
-----

Analysis of Deviance Table (Type II tests)

Response: Experiencia
      Df      F Pr(>F)
Grupo    2 0.8692 0.4287
Residuals 33

-----
Linearity between ErrosDirecao (VD) and Experiencia (covariate)
-----

-----
Placebo
-----

Harvey-Collier test

data: modelo
HC = 0.48977, df = 9, p-value = 0.636

-----
PoucoAlcool
-----

Harvey-Collier test

data: modelo
HC = 0.75424, df = 9, p-value = 0.47
```



```

-----
MuitoAlcool
-----

Harvey-Collier test

data:  modelo
HC = 0.00074081, df = 9, p-value = 0.9994

-----
Homogenous slope of regressions
-----

Analysis of Deviance Table (Type II tests)

Response: ErrosDirecao
      Df      F    Pr(>F)
Grupo    2 19.5215 3.716e-06 ***
Experiencia  1 121.0825 4.711e-12 ***
Grupo:Experiencia  2  0.1583  0.8543
Residuals   30
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Os dois testes das suposições apresentados nesta saída são:

- Dissociação entre Grupo (factor) and Experiência (covariável)

A hipótese nula é a ausência de associação entre os níveis do fator e os níveis de experiência dos voluntários. Isto é feito por uma ANOVA independente com:

```

modelo <- lm(Experiencia ~ Grupo, data=Dados)
print(Anova <- car::Anova(modelo, white.adjust=TRUE))

```

obtendo-se:

O valor $p = 0.4287$ não rejeita H_0 e, portanto, assumimos dissociação entre o fator e a covariável.

- Linearidade entre ErrosDirecao (VD) e Experiencia (covariavel) em todos os níveis do Grupo (fator)

As hipóteses nulas para o teste de Harvey-Collier são de linearidade entre as variáveis testadas. São feitas regressões segmentando os dados para cada grupo com:

```

levels <- as.vector(unique(Dados$Grupo))
for (l in levels)
{
  modelo <- lm(ErrosDirecao ~ Experiencia, data=Dados[Dados$Grupo==l,])
  hc <- lmtest::harvtest(modelo) # Harvey-Collier test for linearity
  cat(bartitle(l,4))
  print(hc)
}

```

```

-----
Placebo
-----

Harvey-Collier test

```

```
data: modelo
HC = 0.48977, df = 9, p-value = 0.636
```

```
-----
PoucoAlcool
-----
```

Harvey-Collier test

```
data: modelo
HC = 0.75424, df = 9, p-value = 0.47
```

```
-----
MuitoAlcool
-----
```

Harvey-Collier test

```
data: modelo
HC = 0.00074081, df = 9, p-value = 0.9994
```

obtendo-se:

Para os três níveis do fator não rejeitou-se a hipótese nula.

- Homogeneidade da inclinação das regressões lineares simples

A hipótese nula é o paralelismo das três regressões feitas, em cada nível do fator, entre a experiência (covariável) e o número de erros cometidos no simulador (VD). Isto é feito por um GLM que considera o fator, a covariável e a interação entre fator e covariável com:

```
modelo <- lm(ErrosDirecao ~ Grupo + Experiencia + Grupo:Experiencia,
             data=Dados)
suppressMessages(
  print(Anova <- car::Anova(modelo, white.adjust=TRUE))
)
```

obtendo-se:

A inclusão da interação Grupo:Experiencia é um artifício para o teste do paralelismo das retas. O valor desta interação, $p = 0.8543$ é o único relevante para a decisão estatística. Neste caso, não rejeitamos H_0 e, portanto, as três regressões lineares simples são assumidas como retas paralelas.

Teste do efeito do fator

A ANCOVA propriamente dita tem hipótese nula de ausência de efeito do fator sobre a VD:

$$H_0 : \mu'_1 = \mu'_2 = \dots = \mu'_k$$

onde k é o número de níveis do fator e μ'_i são as médias populacionais ajustadas da VD pela covariável.

Uma vez que as três regressões são paralelas, detectamos o efeito do fator (os níveis de álcool) sobre a VD (o número de erros cometidos), controlando-se para a covariável (experiência em dirigir dos voluntários de cada nível do fator) caso as regressões tenham diferentes interceptos.

Implementado em demo_ANCOVA_04ANCOVA.R:

```

-----
Homogenous intercepts of regressions
(assuming homogeneous slopes)
- Factor Effect Test -
-----

```

Coefficient covariances computed by hccm()

Analysis of Deviance Table (Type II tests)

Response: ErrosDirecao

	Df	F	Pr(>F)
Grupo	2	14.496	3.295e-05 ***
Experiencia	1	46.304	1.080e-07 ***
Residuals	32		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Por isso, a ANCOVA propriamente dita é obtida por

```

ancova.fit <- lm(ErrosDirecao~Grupo+Experiencia,
                 data=Dados)
print(Anova <- car::Anova(ancova.fit, white.adjust=TRUE))

```

Note que a sintaxe é similar à anterior, exceto pela remoção da interação Grupo:Experiencia. Computa-se os erros de direção em função do fator (Grupo) e da covariável (Experiencia); é uma ANOVA da ANCOVA.

obtendo-se

O valor $p = 3.295 \cdot 10^{-05}$ relevante aqui é aquele associado com o fator (Grupo), rejeitando-se a hipótese nula da ANCOVA que foi enunciado acima, portanto considera-se que a ingestão de álcool afeta o número de erros cometidos na direção.

Na sequência exibimos, também, a regressão associada à esta ANCOVA com

```
print(summary(ancova.fit))
```

Call:

```
lm(formula = ErrosDirecao ~ Grupo + Experiencia, data = Dados)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.5989	-0.9540	-0.0951	0.8361	4.0463

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	9.08210	0.66541	13.649	6.85e-15 ***
GrupoPoucoAlcool	1.44533	0.70939	2.037	0.0499 *
GrupoMuitoAlcool	4.65651	0.69322	6.717	1.38e-07 ***
Experiencia	-0.26165	0.03629	-7.210	3.45e-08 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.696 on 32 degrees of freedom

Multiple R-squared: 0.762, Adjusted R-squared: 0.7397

F-statistic: 34.15 on 3 and 32 DF, p-value: 4.316e-10

O valor que aparece na coluna Estimate na linha da covariável Experiencia é a inclinação comum estimada para as três retas de regressão, que será utilizada adiante.

As últimas duas linhas da saída da regressão da ANCOVA mostram o efeito global ($\eta^2 = R^2 = 0.762$) e o valor $p = 4.316 \cdot 10^{-10}$ que rejeita a hipótese nula da ausência de modelo; portanto, a ANCOVA proposta é um modelo válido para esta análise.

Tamanhos de efeito

Como descrito anteriormente, computa-se o tamanho de efeito global e os tamanhos de efeito parciais, implementadas em `demo_ANCOVA_05tamefeito.R`:

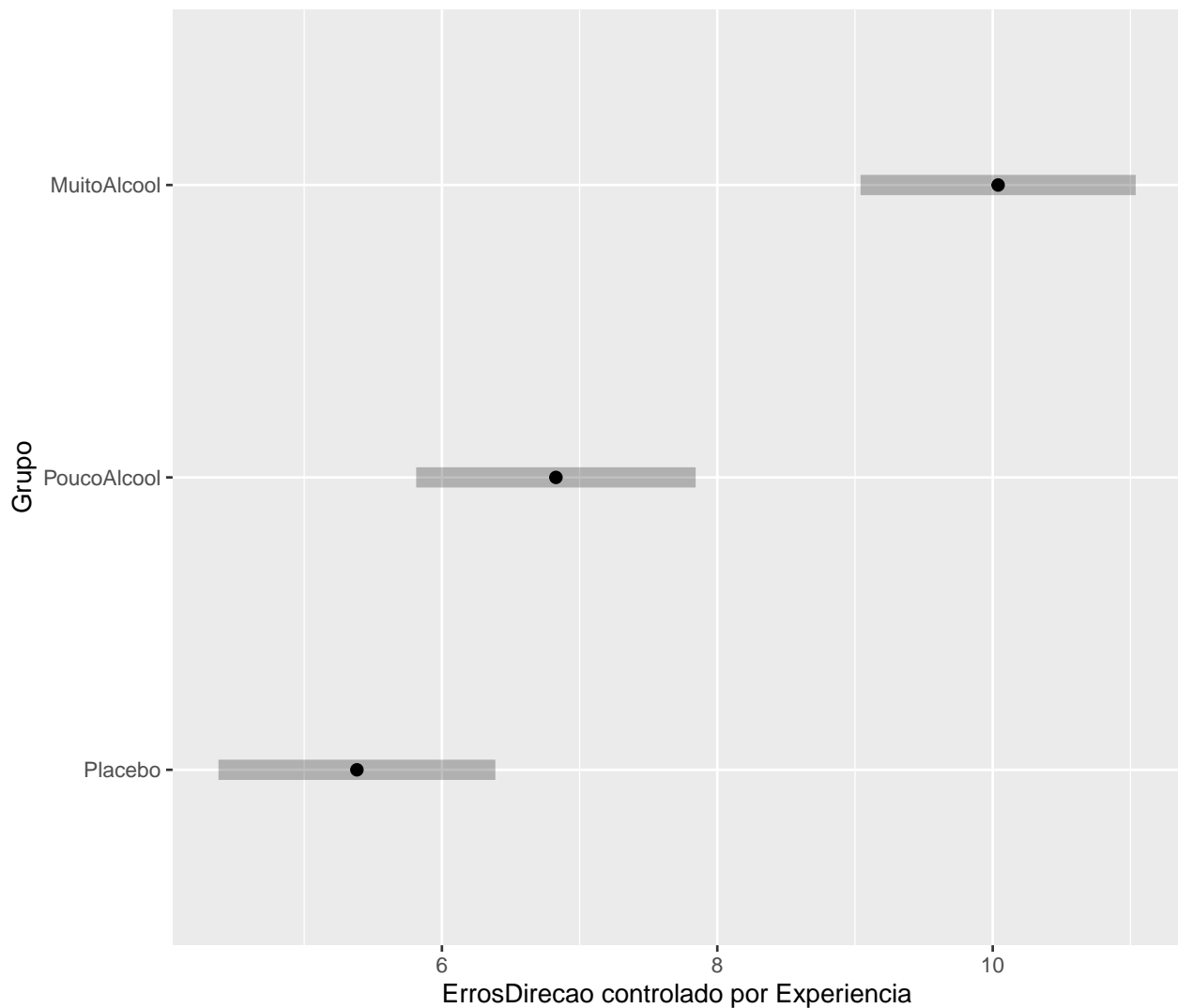
```
-----  
Effect size analysis  
-----  
  
-----  
Omnibus  
-----  
  
Eta^2 = 0.7619839  
large effect size  
  
-----  
Partials  
-----  
  
Partial eta^2 classification  
Grupo      0.5969442      large  
Experiencia 0.6189608      large
```

Médias marginais estimadas ajustadas

O próximo item da saída são as médias marginais e erros-padrão ajustados. Estes são os valores estimados depois que os grupos foram estatisticamente equalizados pela covariável (i.e., os valores estimados para os erros de direção associados ao álcool sem o efeito da experiência). Por isso, são diferentes dos valores observados.

Este trecho está implementado em `demo_ANCOVA_06emmeans.R`:

```
-----  
Adjusted estimated marginal means  
-----  
  
Grupo      emmean    SE df lower.CL upper.CL  
Placebo      5.38 0.494 32     4.38     6.39  
PoucoAlcool  6.83 0.498 32     5.81     7.84  
MuitoAlcool 10.04 0.490 32     9.04    11.04  
  
Confidence level used: 0.95
```



Esta correção dos valores é a partir da regressão linear associada à ANCOVA:

```
suppressMessages(library(readxl))

Dados <- readxl::read_excel("AlcoholExperienceDrivingErrors.xlsx")
Dados$Grupo <- factor(Dados$Grupo,
                      levels=c("Placebo", "PoucoAlcool", "MuitoAlcool"))
grp <- unique(Dados$Grupo)
ngrp <- length(unique(Dados$Grupo))

ancova.fit <- lm(ErrosDirecao~Grupo+Experiencia,
                 data=Dados)
print(reg <- summary(ancova.fit))
```

Call:
lm(formula = ErrosDirecao ~ Grupo + Experiencia, data = Dados)

Residuals:

Min	1Q	Median	3Q	Max
-3.5989	-0.9540	-0.0951	0.8361	4.0463

```

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      9.08210    0.66541  13.649 6.85e-15 ***
GrupoPoucoAlcool  1.44533    0.70939   2.037  0.0499 *
GrupoMuitoAlcool  4.65651    0.69322   6.717 1.38e-07 ***
Experiencia      -0.26165    0.03629  -7.210 3.45e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.696 on 32 degrees of freedom
Multiple R-squared:  0.762, Adjusted R-squared:  0.7397
F-statistic: 34.15 on 3 and 32 DF,  p-value: 4.316e-10

```

Sendo uma regressão, a coluna Estimate contém o intercepto associado ao primeiro nível do fator (9.0821029), os interceptos dos outros níveis relativos ao primeiro nível (1.4453283 e 4.6565087), e a inclinação da reta comum a todos os níveis (-0.2616459). O objetivo é obter três retas com a mesma inclinação.

Os valores absolutos dos interceptos são dados pela soma dos valores relativos sobre o intercepto do primeiro nível:

```

# interceptos ajustados
adj_intercept <- rep(NA, ngrp)
adj_intercept[1] <- reg$coefficients[1,1]
for (g in 2:ngrp)
{
  adj_intercept[g] <- adj_intercept[1] + reg$coefficients[g,1]
}
cat("\n")
cat("Intercepts:\n")
cat("\t", as.character(grp[1]), " = ", adj_intercept[1], "\n", sep="")
for (g in 2:ngrp)
{
  cat("\t", as.character(grp[g]), " = ", adj_intercept[1], " + ", reg$coefficients[g,1], " = ", adj_intercept[g], "\n", sep="")
}

```

Intercepts:

```

Placebo = 9.082103
PoucoAlcool = 9.082103 + 1.445328 = 10.52743
MuitoAlcool = 9.082103 + 4.656509 = 13.73861

```

A inclinação comum às três retas é dada por

```

adj_slope <- reg$coefficients[ngrp+1,1]
cat("Common slope = ", adj_slope, "\n", sep="")

```

Common slope = -0.2616459

e a média comum da covariável é a média das médias da experiência dos voluntários em cada um dos grupos:

```

tmp <- aggregate(Experiencia~Grupo, data=Dados, FUN=mean)
print(tmp)

```

```

      Grupo Experiencia
1  Placebo    12.41667
2 PoucoAlcool  16.66667
3 MuitoAlcool  13.33333

```

```
adj_cov <- mean(tmp$Experiencia)
cat("Common Covariate Mean = ",adj_cov,"\n", sep="")
```

```
Common Covariate Mean = 14.13889
```

Finalmente, as médias marginais estimadas ajustadas correspondem aos valores do número de erros de direção para a média comum da covariável estimados pelas três regressões obtidas com os interceptos ajustados e inclinação comum:

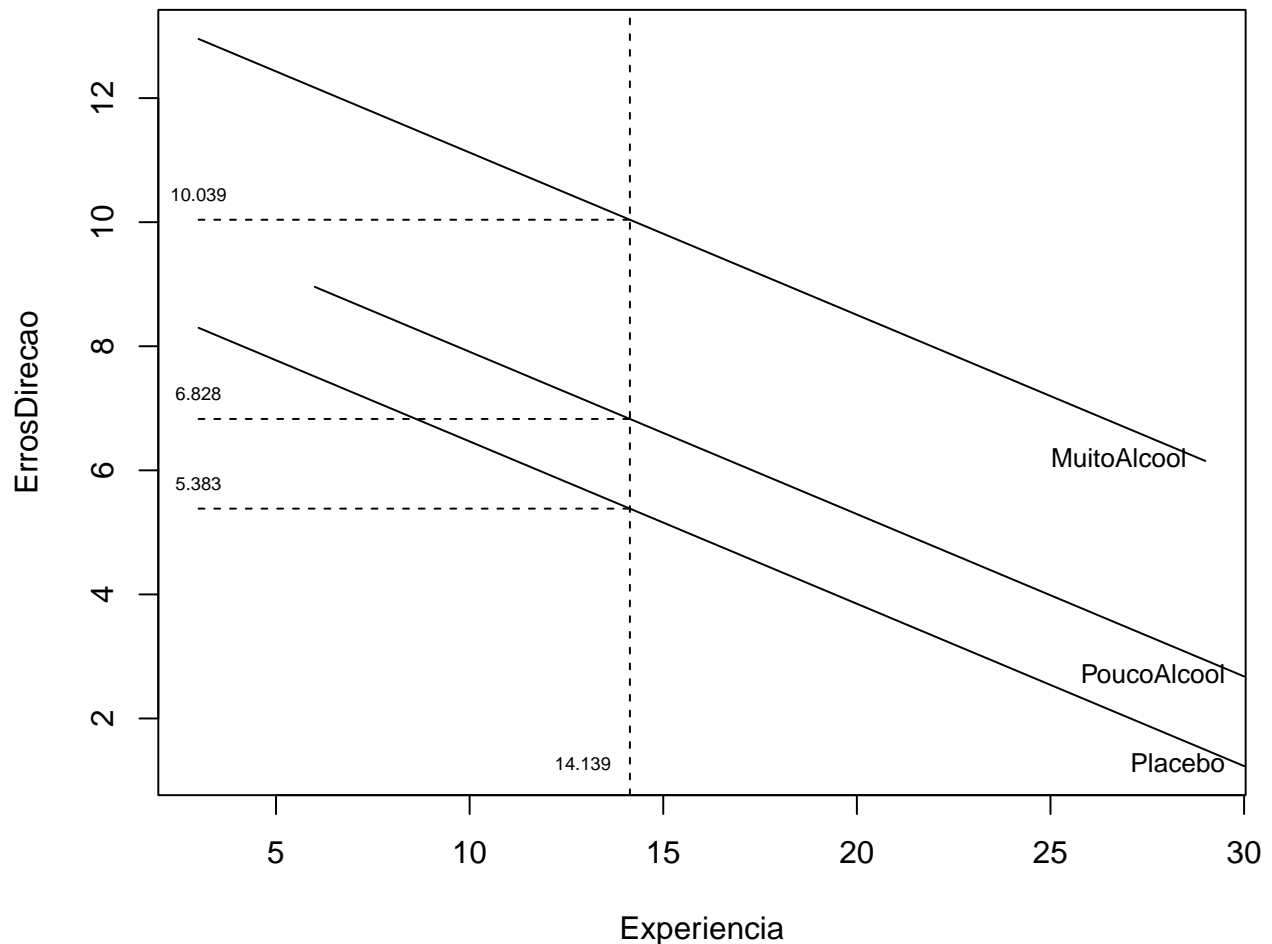
```
adj_mm <- rep(NA,ngroup)
for (g in 1:ngroup)
{
  adj_mm[g] <- adj_intercept[g] + adj_slope*adj_cov
}
cat("Adjusted marginal means:\n")
for (g in 1:ngroup)
{
  cat("\t",as.character(grp[g])," = ",adj_intercept[g]," + ",adj_slope," * ",adj_cov," = ",adj_mm[g],"\n")
}
```

```
Adjusted marginal means:
```

```
Placebo = 9.082103 + -0.2616459 * 14.13889 = 5.382721
PoucoAlcool = 10.52743 + -0.2616459 * 14.13889 = 6.828049
MuitoAlcool = 13.73861 + -0.2616459 * 14.13889 = 10.03923
```

Uma visão gráfica destas médias marginais ajustadas é:

Adjusted Marginal Means



onde 14.139 é a média comum da covariável e os valores das médias marginais correspondem aos que foram obtidos com a função `emmeans::emmeans()`

```
suppressMessages(library(emmeans))
```

```
alfa <- 0.05
EMM <- emmeans::emmeans(ancova.fit,
                          pairwise~"Grupo",
                          adjust="tukey",
                          level=1-alfa)
print(EMM$emmeans)
```

Grupo	emmean	SE	df	lower.CL	upper.CL
Placebo	5.38	0.494	32	4.38	6.39
PoucoAlcool	6.83	0.498	32	5.81	7.84
MuitoAlcool	10.04	0.490	32	9.04	11.04

Confidence level used: 0.95

No mesmo processo de cálculo das médias marginais, foram computados também erros padrão ajustados (SE, *standard errors*), que costumam ter valores quase iguais para cada nível do fator. O cálculo para estes erros padrão é um pouco mais trabalhoso (Liu, 2011; Crawley, 2012). Basta dizer aqui que estes valores são os

utilizados para estimar o intervalo de confiança que aparece no gráfico correspondente.



O código completo para calcular as médias marginais estimadas ajustadas está em `demo_AdjustedMarginalMeans.R`.

Testes *post hoc*

Resta apenas localizar as diferenças dos números de erros entre os diferentes níveis do fator. Usamos os contrastes de Tukey da mesma forma que fizemos com ANOVA nos capítulos anteriores. Este trecho está implementado em `demo_ANCOVA_07posthoc.R`:

```
-----  
Post hoc tests: Grupo (factor) adjusted by Experiencia (covariate)  
-----
```

Legend:

```
A ... Placebo  
B ... PoucoAlcool  
C ... MuitoAlcool
```

Simultaneous Confidence Intervals

Multiple Comparisons of Means: Tukey Contrasts

```
Fit: lm(formula = ErrosDirecao ~ GrupoShort + Experiencia, data = Dados)
```

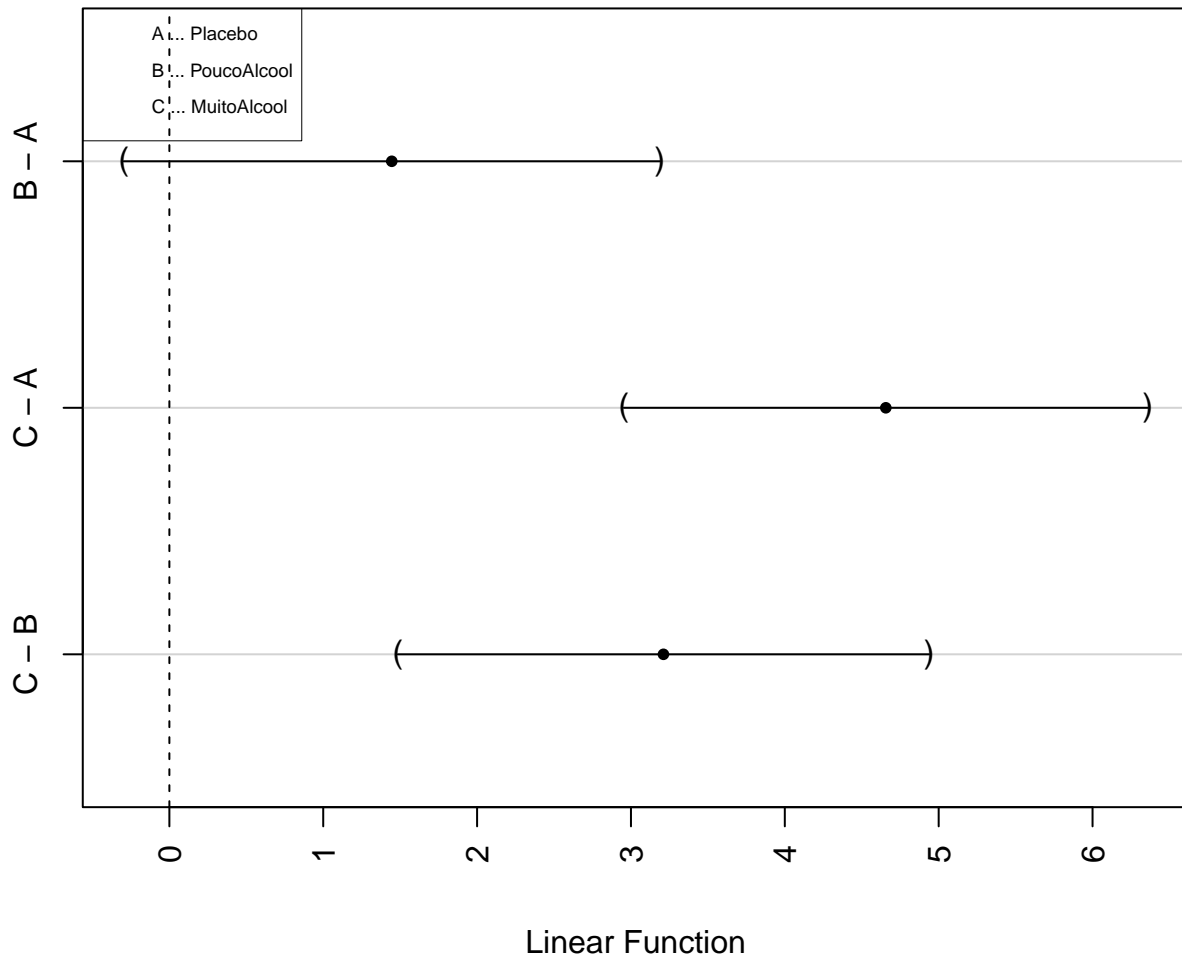
```
Quantile = 2.4557
```

```
95% family-wise confidence level
```

Linear Hypotheses:

	Estimate	lwr	upr
B - A == 0	1.4453	-0.2967	3.1874
C - A == 0	4.6565	2.9542	6.3588
C - B == 0	3.2112	1.4851	4.9373

Tukey contrasts: Grupo 95% family-wise CI



Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Fit: `lm(formula = ErrosDirecao ~ GrupoShort + Experiencia, data = Dados)`

Linear Hypotheses:

	Estimate	Std. Error	t value	Pr(> t)
B - A == 0	1.4453	0.7094	2.037	0.149841
C - A == 0	4.6565	0.6932	6.717	4.15e-07 ***
C - B == 0	3.2112	0.7029	4.568	0.000208 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- bonferroni method)

A	B	C
"a"	"a"	"b"

Concluimos, portanto que os níveis A (Placebo) e B (PoucoAlcool) são estatisticamente iguais, mas o nível

C (MuitoAlcool) é o motivo da rejeição da hipótese nula e, portanto, Álcool tem efeito no número de erros cometidos no simulador de direção.

@@

ANCOVA mista

Em outra situação, os mesmos indivíduos são avaliados antes e depois de uma intervenção (intraparticipante), mas são divididos em grupos que recebem tratamentos diferentes (entre participantes). O interesse é avaliar o efeito das tratamentos.

Segundo BORM et al., 2007:

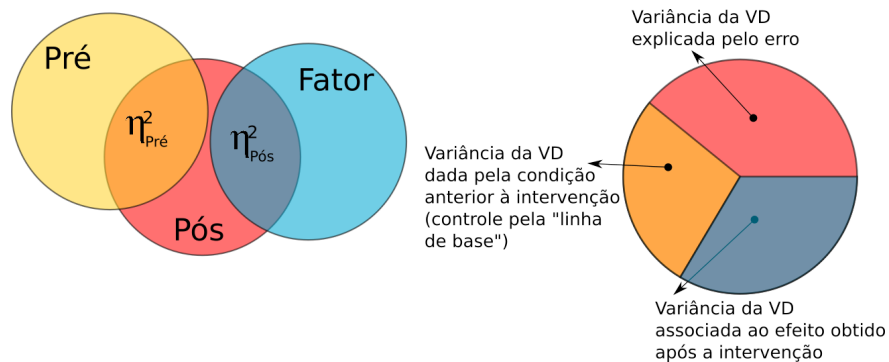
“An advantage of the use of ANCOVA is that it adjusts for baseline differences between the treatment groups. ANCOVA also has more statistical power than the t-test, so sample size requirements are lower.”

Exemplo

Dez pacientes idosos foram submetidos a uma cirurgia de grande porte com anestesia geral. Divididos em dois grupos, cada um tratado com uma droga diferente: uma droga tradicional (padrão) e uma nova proposta terapêutica (nova). Um teste cognitivo foi aplicado alguns dias antes da cirurgia e no período pós-operatório. Espera-se que as drogas mitiguem delírio pós-operatório (DPO) e disfunção cognitiva pós-operatória (DCPO). Este estudo é duplo cego para a comparação de duas drogas, de forma que serão denominadas de droga A e droga B, sem que o analista saiba qual delas corresponde à droga padrão e qual corresponde à droga nova. Os dados obtidos estão em PrePos_2CondIndep.xlsx, organizados no formato *long*.

“A anestesia geral age predominantemente no sistema nervoso central, repercutindo também, em todos os aparelhos e sistemas do organismo. Seu mecanismo intrínseco de ação ainda não é completamente conhecido e por isso a possibilidade de algum prejuízo temporário ou permanente na cognição e na memória sempre foi alvo de considerações. Há uma especial preocupação quanto aos idosos, por apresentarem maior susceptibilidade às alterações da homeostasia e do meio ambiente. A cirurgia e a anestesia exercem comparativamente efeitos adversos cerebrais mais acentuados nos idosos do que nos jovens, manifestado pela maior prevalência de delírio pós-operatório e disfunção cognitiva. O delírio pós-operatório (DPO) e a disfunção cognitiva pós-operatória (DCPO) atrasam a reabilitação, e são associadas com o aumento na morbidade e na mortalidade de pacientes idosos. Embora seja difícil estabelecer metodologicamente qualquer correlação entre DPO e DCPO, um recente estudo sugere que ambas podem representar uma trajetória de insuficiência cognitiva pós-operatória, talvez como uma progressão de não reconhecida insuficiência cognitiva leve pré-operatória. [...] A incidência de DCPO é elevada (40%) em pacientes com mais de 65 anos, submetidos a cirurgias de médio e grande porte e, em cirurgias cardíacas, esta porcentagem pode ser maior que 60%. [...] Aproximadamente 10% dos pacientes cirúrgicos idosos desenvolvem DPO, aumentando para 30-65% após certos tipos de cirurgia, tais como cirurgia cardíaca, cirurgias de emergência e fratura de quadril.” (Sakae et al., 2016)

Esta ANCOVA é similar à independente, mas o papel da covariável é feito pela medida anterior à intervenção (Pre), que funciona como controle da linha de base.



Quem faz o papel da variável de desfecho é a medida posterior à intervenção (Pos). O fator é a droga, e neste exemplo existem dois níveis (droga A e B); são dois grupos de pacientes independentes, mas cada um deles é medido duas vezes. Neste sentido o delineamento é misto: entre-participantes e intra-participantes.

Esta análise foi implementada em `demo_ANCOVA_prepos.R`. O código é praticamente o mesmo que o anterior, com pequenos ajustes por causa da estatística descritiva e pelo uso dos formatos *long* e *wide* de acordo com a conveniência das funções utilizadas.

Obtém-se a seguinte saída:

```

----
Data
----
  Paciente Grupo  Momento  Escore
1         droga A Pre      150
2         droga A Pre      130
3         droga A Pre      125
4         droga A Pre      152
5         droga A Pre      160
6         droga B Pre      174
7         droga B Pre      110
8         droga B Pre      180
9         droga B Pre      145
10        droga B Pre      140
1         droga A Pos       51
2         droga A Pos       50
3         droga A Pos       40
4         droga A Pos       45
5         droga A Pos       60
6         droga B Pos       75
7         droga B Pos       41
8         droga B Pos       80
9         droga B Pos       60
10        droga B Pos       55

-----
Descriptive Statistics
-----
, , Momento = Pre

      Grupo
Escore droga A droga B
      40      0      0

```

```

41      0      0
45      0      0
50      0      0
51      0      0
55      0      0
60      0      0
75      0      0
80      0      0
110     0      1
125     1      0
130     1      0
140     0      1
145     0      1
150     1      0
152     1      0
160     1      0
174     0      1
180     0      1

, , Momento = Pos

      Grupo
Escore droga A droga B
40      1      0
41      0      1
45      1      0
50      1      0
51      1      0
55      0      1
60      1      1
75      0      1
80      0      1
110     0      0
125     0      0
130     0      0
140     0      0
145     0      0
150     0      0
152     0      0
160     0      0
174     0      0
180     0      0

      Escore 40 41 45 50 51 55 60 75 80 110 125 130 140 145 150 152 160 174 180
Grupo  Momento
droga A Pre      0  0  0  0  0  0  0  0  0  0  0  1  1  0  0  1  1  1  0  0
        Pos      1  0  1  1  1  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0
droga B Pre      0  0  0  0  0  0  0  0  0  0  1  0  0  1  1  0  0  0  1  1
        Pos      0  1  0  0  0  1  1  1  1  0  0  0  0  0  0  0  0  0  0  0

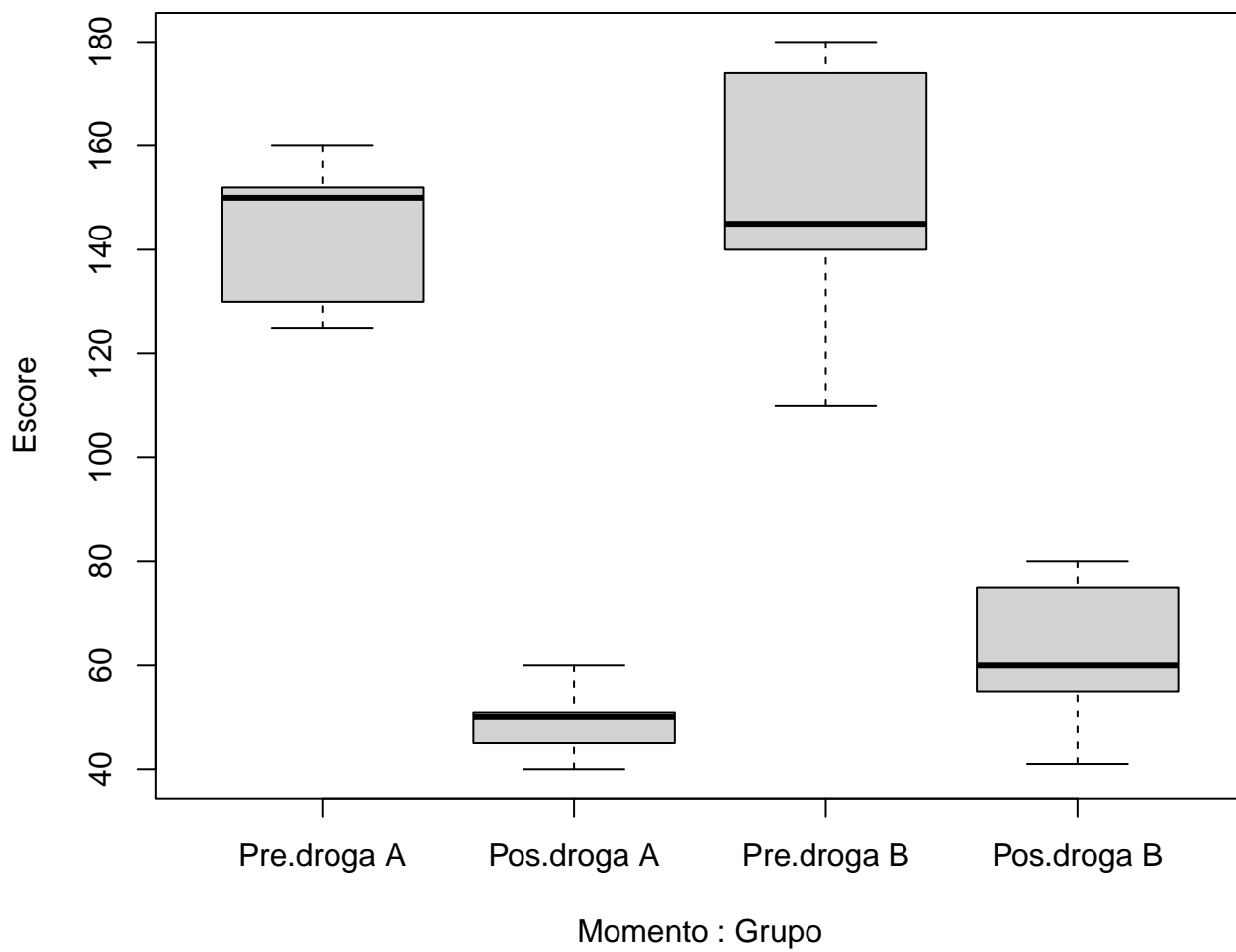
Means Pre:
  Grupo[Dados.long$Momento == "Pre"] Escore[Dados.long$Momento == "Pre"]
1                                droga A                                143.4

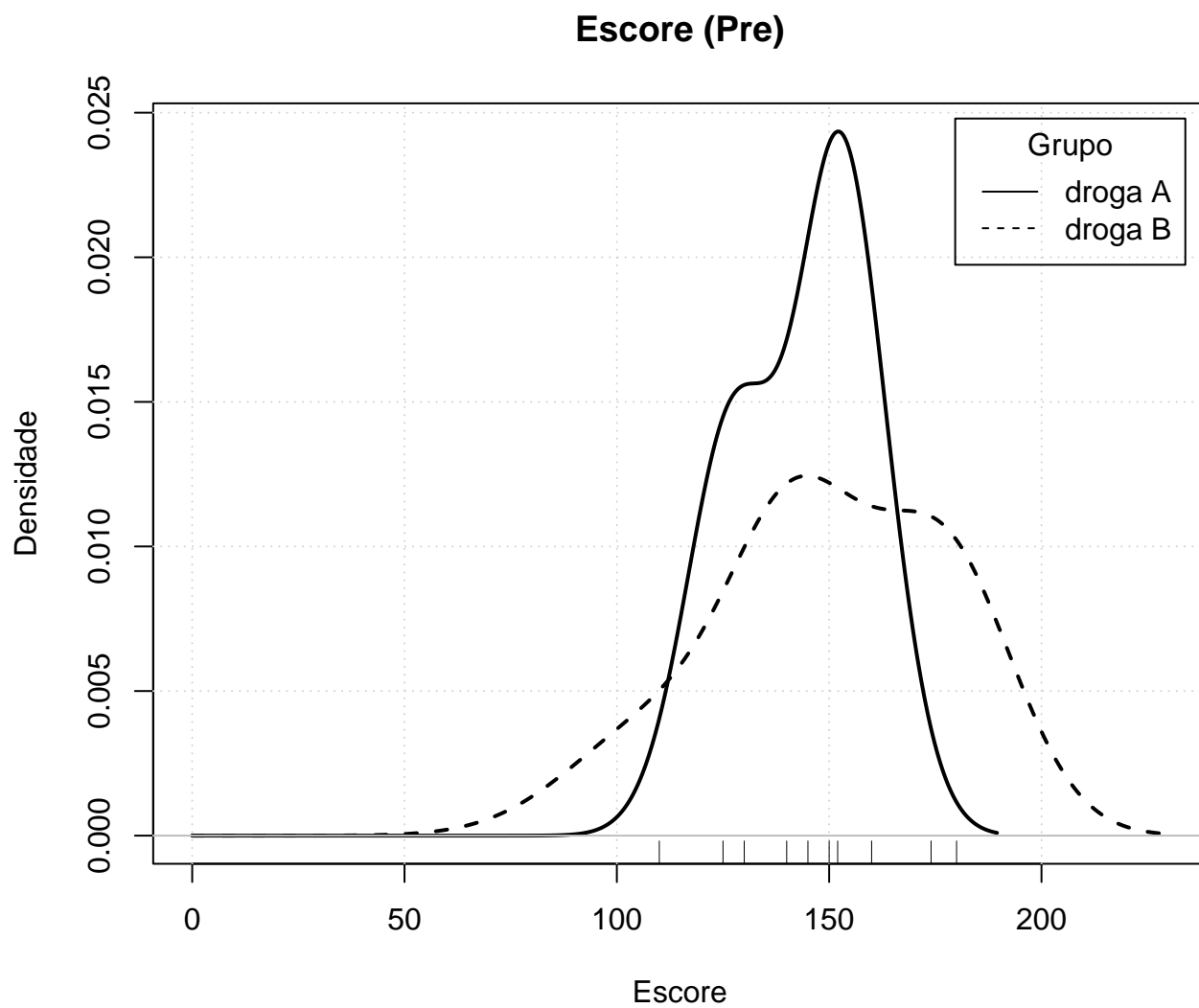
```

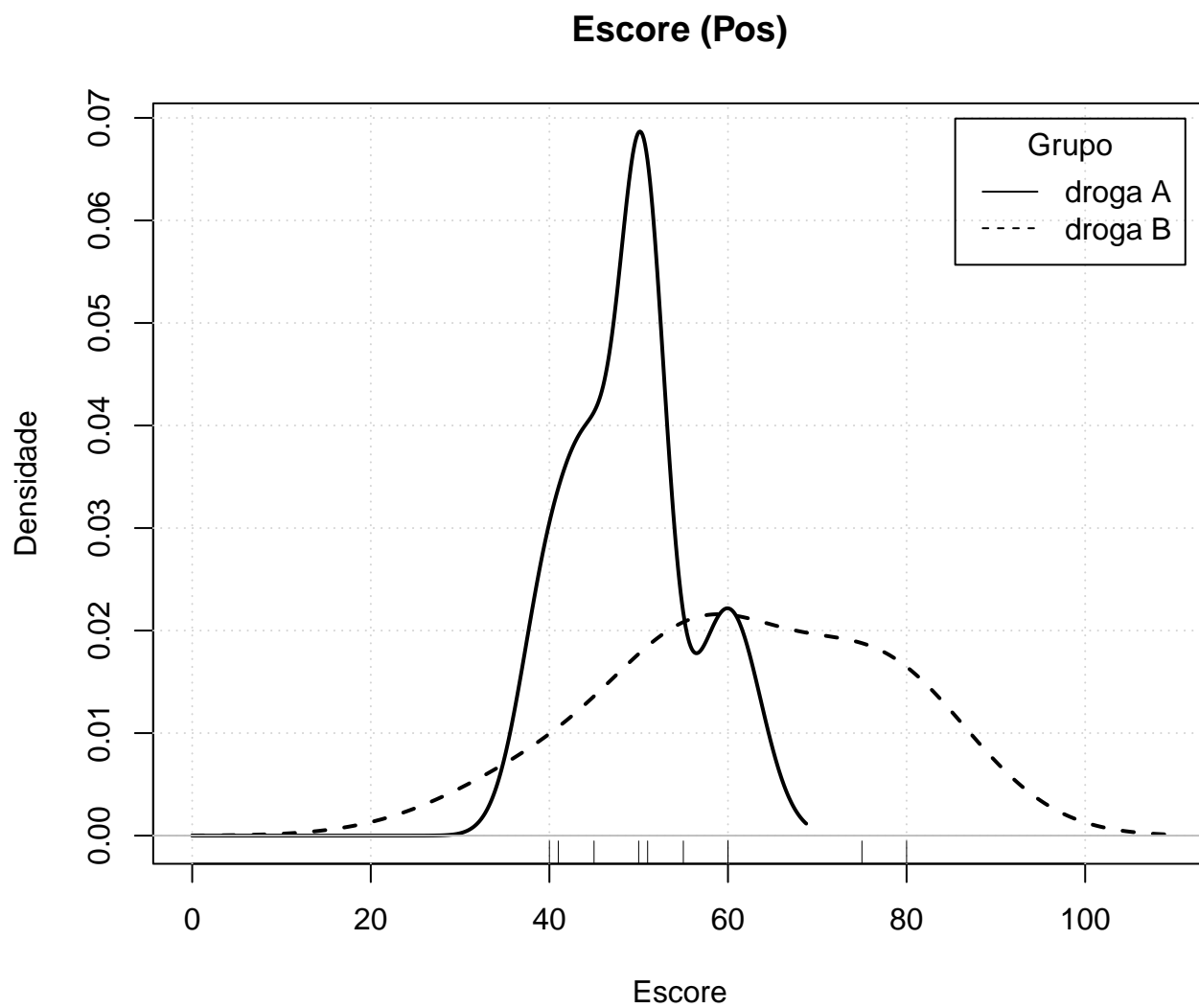
2	droga B	149.8
---	---------	-------

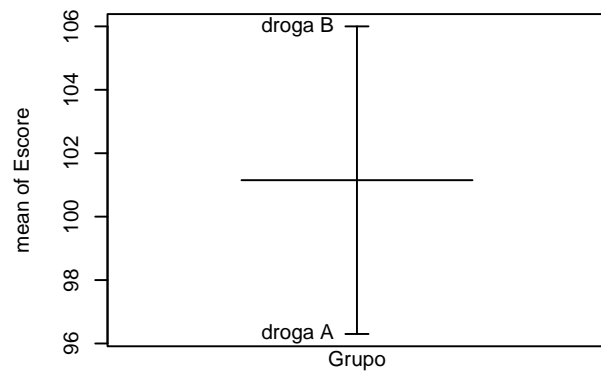
Means Pos:

Grupo[Dados.long\$Momento == "Pos"] Escore[Dados.long\$Momento == "Pos"]		
1	droga A	49.2
2	droga B	62.2

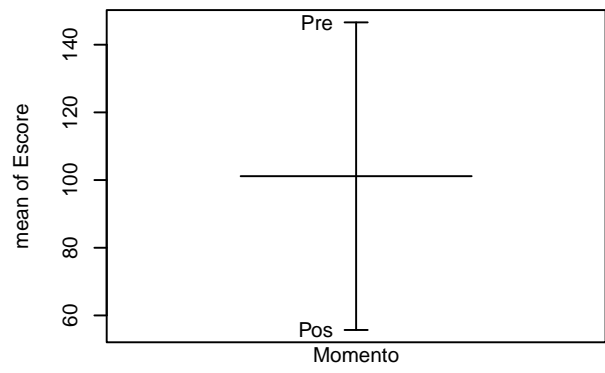




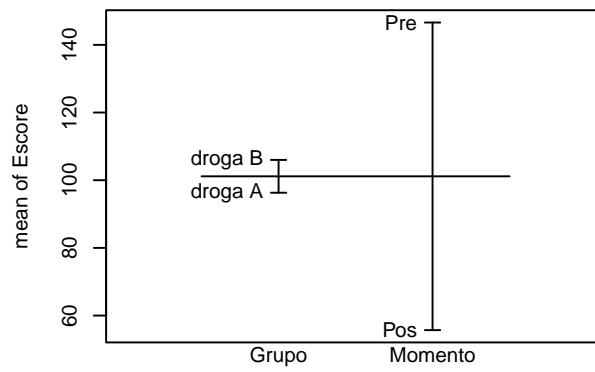




Factors

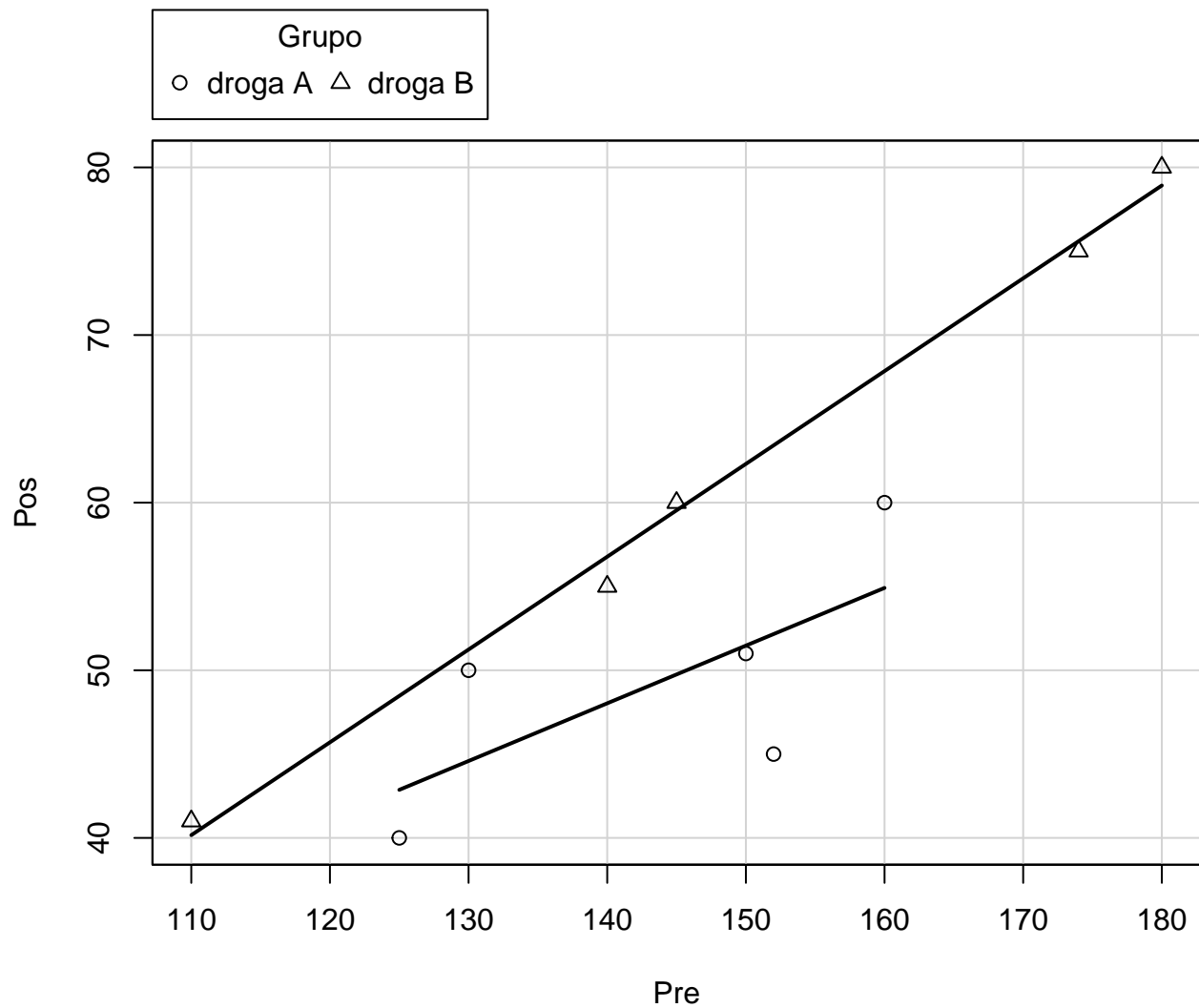


Factors



Factors

Simple linear regressions



`geom_smooth()` using formula 'y ~ x'

```
-----
droga A
-----
```

Call:

```
lm(formula = Pos ~ Pre, data = dt_tmp)
```

Residuals:

1	2	3	4	5
-0.4715	5.4118	-2.8674	-7.1598	5.0869

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.1528	29.5259	-0.005	0.996
Pre	0.3442	0.2050	1.679	0.192

Residual standard error: 6.188 on 3 degrees of freedom

Multiple R-squared: 0.4844, Adjusted R-squared: 0.3126

F-statistic: 2.819 on 1 and 3 DF, p-value: 0.1918

droga B

Call:

lm(formula = Pos ~ Pre, data = dt_tmp)

Residuals:

1	2	3	4	5
-0.5989	0.8362	1.0790	0.4576	-1.7740

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-20.7404	3.6618	-5.664	0.010900 *
Pre	0.5537	0.0241	22.971	0.000181 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.364 on 3 degrees of freedom

Multiple R-squared: 0.9943, Adjusted R-squared: 0.9925

F-statistic: 527.7 on 1 and 3 DF, p-value: 0.0001807

ANCOVA

Assumptions

Dissociation between Grupo (factor) and VD Pre (covariate)

Coefficient covariances computed by hccm()

Analysis of Deviance Table (Type II tests)

Response: Pre

	Df	F	Pr(>F)
Grupo	1	0.1594	0.7002
Residuals	8		

Linearity between Escore pos (VD) and Escore pre (covariate)

droga A

Harvey-Collier test

data: modelo

HC = 0.30952, df = 2, p-value = 0.7862

droga B

Harvey-Collier test

data: modelo

HC = 0.29829, df = 2, p-value = 0.7936

Homogenous slope of regressions

Coefficient covariances computed by hccm()

Analysis of Deviance Table (Type II tests)

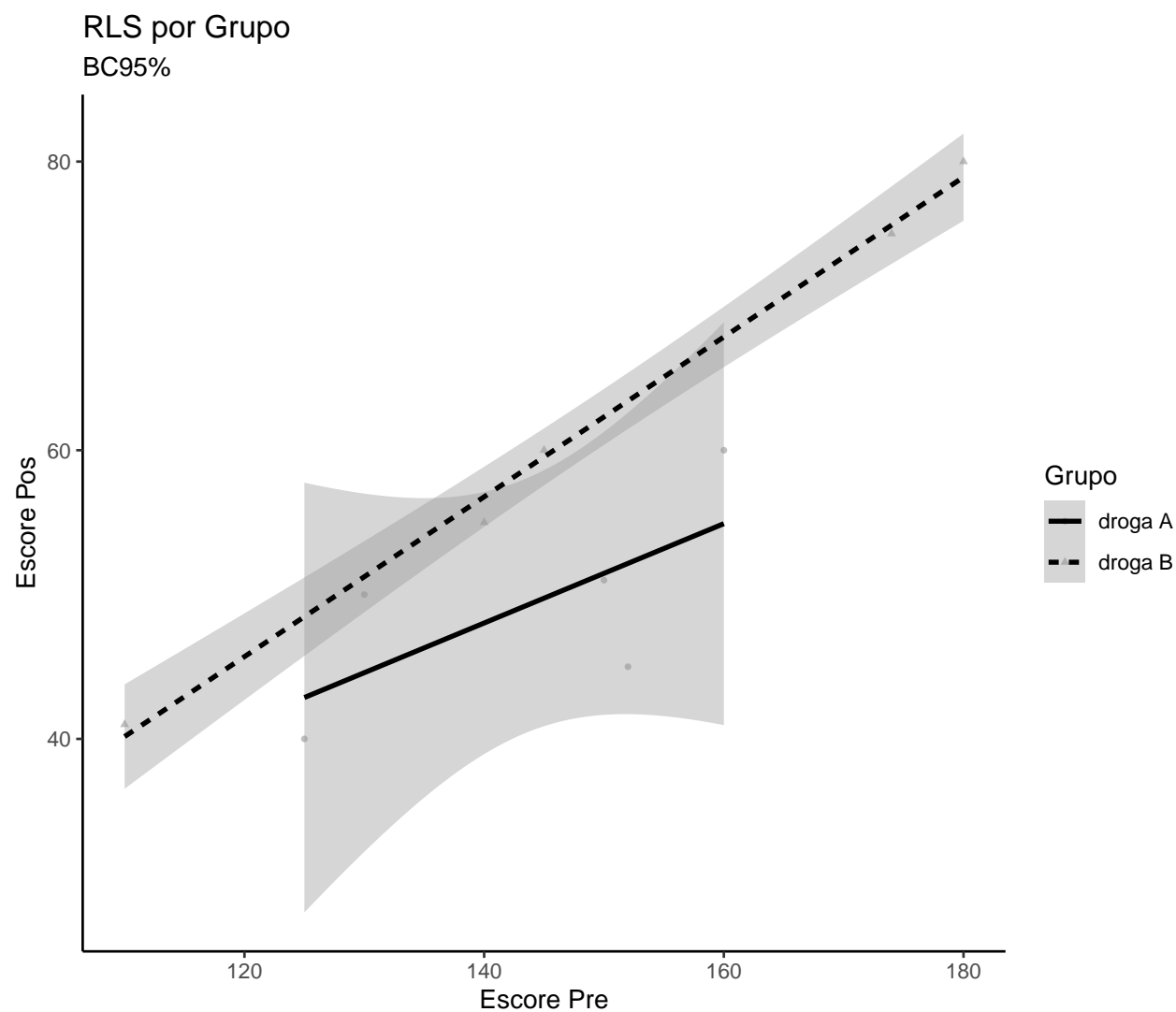
Response: Pos

	Df	F	Pr(>F)
Grupo	1	5.9729	0.0502 .
Pre	1	186.0164	9.648e-06 ***
Grupo:Pre	1	0.5427	0.4891
Residuals	6		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Homogenous intercepts of regressions
(assuming homogeneous slopes)
- Factor Effect Test -

Coefficient covariances computed by hccm()



Analysis of Deviance Table (Type II tests)

Response: Pos

	Df	F	Pr(>F)
Grupo	1	8.9356	0.02025 *
Pre	1	75.7297	5.306e-05 ***
Residuals	7		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Call:

```
lm(formula = Pos ~ Grupo + Pre, data = Dados)
```

Residuals:

Min	1Q	Median	3Q	Max
-8.5623	-1.4138	0.1841	1.9159	7.5971

Coefficients:

Estimate	Std. Error	t value	Pr(> t)
----------	------------	---------	----------

```

(Intercept)  -23.53923    10.61245   -2.218 0.062053 .
Grupodroga B   9.75362     2.97954    3.274 0.013608 *
Pre           0.50725     0.07257    6.990 0.000213 ***

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.653 on 7 degrees of freedom
Multiple R-squared: 0.9071, Adjusted R-squared: 0.8806
F-statistic: 34.18 on 2 and 7 DF, p-value: 0.0002442

Effect size analysis

Omnibus

Eta² = 0.9071233
large effect size

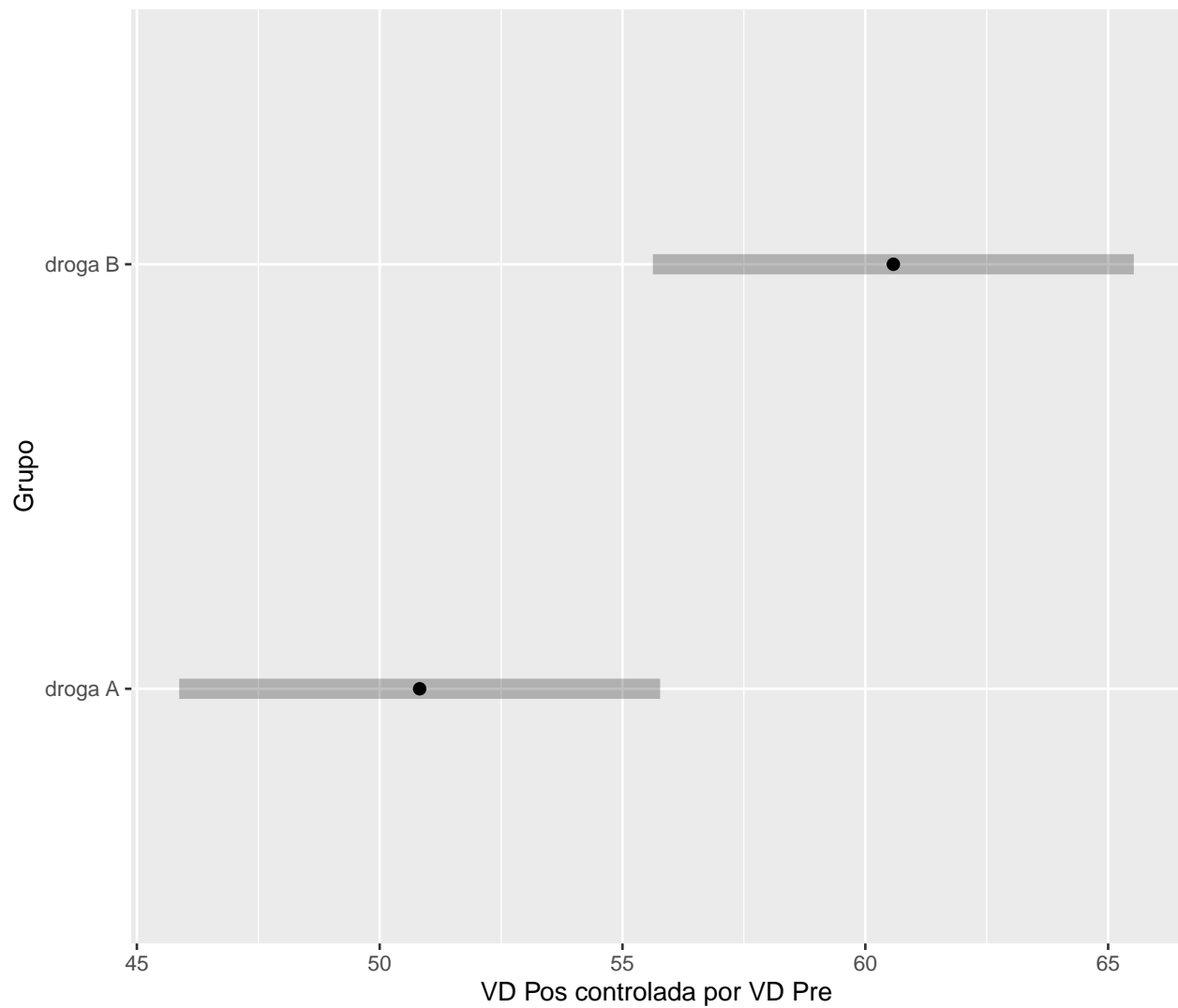
Partial

Partial eta² classification
Grupo 0.6048774 large
Pre 0.8746825 large

Adjusted estimated marginal means

Grupo	emmean	SE	df	lower.CL	upper.CL
droga A	50.8	2.09	7	45.9	55.8
droga B	60.6	2.09	7	55.6	65.5

Confidence level used: 0.95



 Post hoc tests: Grupo (factor) adjusted by VD Pre (covariate)

Legend:

A ... droga A

B ... droga B

Simultaneous Confidence Intervals

Multiple Comparisons of Means: Tukey Contrasts

Fit: `lm(formula = Pos ~ GrupoShort + Pre, data = Dados)`

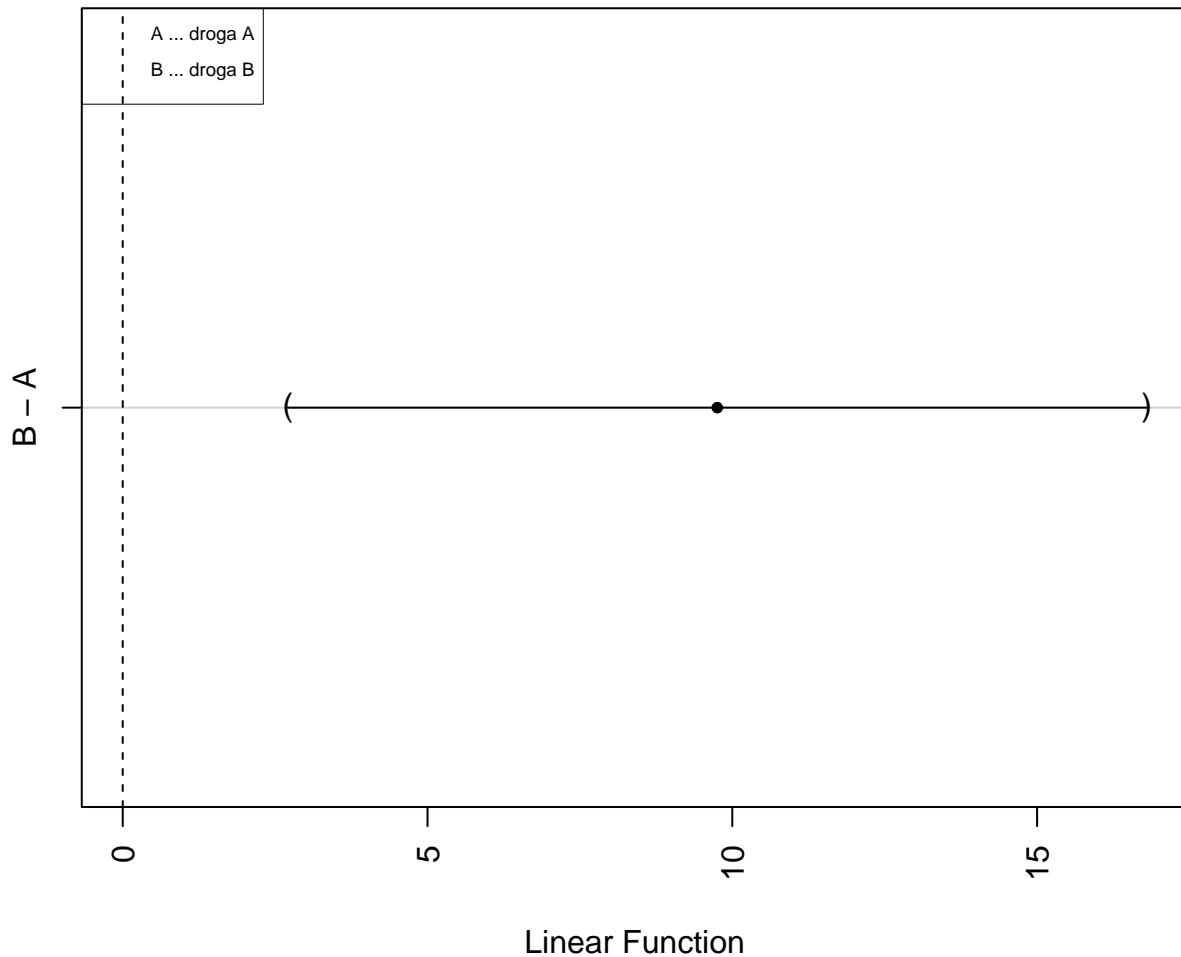
Quantile = 2.3646

95% family-wise confidence level

Linear Hypotheses:

	Estimate	lwr	upr
B - A == 0	9.7536	2.7081	16.7991

Tukey contrasts: Grupo 95% family-wise CI



Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Fit: `lm(formula = Pos ~ GrupoShort + Pre, data = Dados)`

Linear Hypotheses:

	Estimate	Std. Error	t value	Pr(> t)
B - A == 0	9.754	2.980	3.274	0.0136 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- bonferroni method)

A B


```
"a" "b"
```

A interpretação da saída é a mesma da ANOVA independente. Neste exemplo, chamamos a atenção para algumas ocorrências:

1. Nas regressões lineares simples, a droga *A* não tem inclinação significativa para $\alpha = 0.05$. No entanto, o teste da homogeneidade das inclinações não rejeita a hipótese nula de paralelismo entre as regressões. Observe o gráfico correspondente: há um segmento de reta que pode ser traçado pertencendo simultaneamente às bandas de confiança de 95% no domínio comum da covariável.
2. Como acontece na ANCOVA independente, existe a saída da ANOVA da ANCOVA (a ANCOVA propriamente dita) e a saída da regressão da ANCOVA.

Na saída da ANOVA da ANCOVA avalia-se o valor p associado ao fator, que neste caso é 0.02025. Portanto, rejeita-se a hipótese nula para $\alpha = 0.05$ e assume-se que as drogas A e B têm efeitos diversos.

Na saída da regressão da ANCOVA, interessa observar as duas últimas linhas:

Aqui testa-se se existe modelo ANCOVA. O valor $p=0.0002442$ rejeita a hipótese nula de ausência de modelo, portanto há modelo. Também é onde encontramos o tamanho de efeito global ($\eta^2 = R^2$).

3. Como há somente dois níveis do fator, não haveria necessidade do teste *post hoc* que meramente deveria concordar com o que obtém da ANCOVA (está disponível neste *Rscript* para o caso do código ser adaptado para situações com mais níveis do fator). No entanto, observe que os valores p não são idênticos (para ANOVA da ANCOVA, $p = 0.02025$; para o teste *post hoc* $p = 0.0136$) porque a ANCOVA foi feita com correção para heterocedasticidade de White e o teste *post hoc* com correção de Tukey.

Planejamento

Segundo BORM et al. (2007):

“The use of ANCOVA may considerably reduce the number of patients required for a trial. [...] ANCOVA on $(1 - \rho^2) \cdot n$ observations has the same power as a t-test on n observations.”

Estes autores propõem uma fórmula simples para dimensionar uma amostra para a ANCOVA para um fator com dois níveis:

```
delta <- .5
rho <- .6
alfa <- .05
poder <- .90
n.porcondicao <- 2*((qnorm(1-alfa/2)+qnorm(poder))*sqrt(1-rho^2)/delta)^2
n.total <- 2*n.porcondicao
cat("\nn total = ",n.total," com ",n.porcondicao," por nivel do fator\n")
```

```
n total = 107.596 com 53.79801 por nivel do fator
```

Neste *Rscript* encontramos o tamanho da amostra para $\alpha = 0.05$ e poder = 0.90. Temos:

- delta que é o d de Cohen populacional (tamanho de efeito populacional)
- rho que a correlação populacional de Pearson entre a VD e a covariável.

Podemos variar o valor do poder para termos ideia do poder do estudo de nosso exemplo, com 10 idosos no total. Encontramos:

```
delta <- .5
rho <- .6
alfa <- .05
poder <- .17
```

```
n.porcondicao <- 2*((qnorm(1-alfa/2)+qnorm(poder))*sqrt(1-rho^2)/delta)^2
n.total <- 2*n.porcondicao
cat("\nn total = ",n.total," com ",n.porcondicao," por nivel do fator\n")
```

```
n total = 10.3591 com 5.179551 por nivel do fator
```

Portanto, o poder *a priori* de um estudo deste tamanho era bem baixo, da ordem de 17%.

Segundo Liu (2011), embora a inclusão da covariável seja geralmente aceita como capaz de reduzir o erro padrão ajustado, consequentemente aumentando o poder da ANCOVA em comparação com a ANOVA (aumentando o valor de η^2), isto nem sempre se verifica. Com o aumento do tamanho da amostra, a probabilidade de haver benefício com a inclusão de uma covariável também aumenta. O autor apresenta a seguinte tabela:

Table 2
Sample size and probability p_1 and p_2

n	ν	$F1$	p_1	$F2$	p_2
5	8	-0.308	0	-0.684	0
6	10	-0.110	0	-0.405	0
7	12	0.088	0.228	-0.154	0
8	14	0.286	0.399	0.080	0.219
9	16	0.484	0.503	0.305	0.412
10	18	0.681	0.580	0.524	0.522
11	20	0.879	0.640	0.738	0.600
12	22	1.077	0.689	0.950	0.660
13	24	1.275	0.730	1.159	0.708
14	26	1.473	0.764	1.366	0.747
15	28	1.670	0.793	1.571	0.780
16	30	1.868	0.818	1.776	0.807
17	32	2.066	0.840	1.980	0.831
18	34	2.264	0.858	2.183	0.851
19	36	2.462	0.875	2.385	0.869
20	38	2.659	0.889	2.587	0.884
21	40	2.857	0.901	2.789	0.897
22	42	3.055	0.912	2.990	0.909
23	44	3.253	0.922	3.191	0.919
24	46	3.451	0.930	3.391	0.928
25	48	3.648	0.938	3.592	0.936
26	50	3.846	0.945	3.792	0.943
27	52	4.044	0.950	3.992	0.949
28	54	4.242	0.956	4.191	0.954

Neste estudo foram supostos dois níveis para o fator. A correlação de Pearson entre a VD e a covariável é 0.3 (intermediária). Há tamanho total da amostra (n) e graus de liberdade correspondentes da ANCOVA (ν). As colunas p_1 e p_2 são probabilidades (calculadas por dois métodos diversos) da ANCOVA ter resultado melhor que a ANOVA nas mesmas condições: p_1 é a probabilidade de reduzir o erro padrão ajustado; p_2 é probabilidade de reduzir o amplitude do intervalo de confiança ajustado. Observe que para uma probabilidade de 90% de haver benefício com a inclusão da covariável necessita-se de $n \geq 22$.

Para quem quiser modelo mais complexo, Crawley (2012) em “Chapter12 - Analysis of covariance” apresenta

um delineamento com dois fatores e uma covariável.

Referências

Xiaofeng Liu (2011) The Effect of a Covariate on Standard Error and Confidence Interval Width. *Communications in Statistics - Theory and Methods* 40:3, 449-456, DOI: 10.1080/03610920903391337

Sakae TM et al. (2016) Efeitos da anestesia geral da cognição do idoso. *Arquivos Catarinenses de Medicina* 45(3): 107-16.

Borm et al. (2007) A simple sample size formula for analysis of covariance in randomized clinical trials. *Journal of Clinical Epidemiology* 60: 1234-8.

Michael J. Crawley (2012) *The R Book*, 2nd Edition. ISBN: 978-0-470-97392-9, 1076 pages.

Dancey CP & Reidy J (2019) *Estatística sem Matemática para Psicologia*, 7^a ed. Porto Alegre: Penso.