Igor Costa

Métodos estatísticos aplicados à psicolinguística experimental

25 de julho de 2021

…to understand what you are seeing,
you need to know something about
how you would approach the
problem by hand…

Howell, 2010:462

1 Modelos lineares mistos

A fim de começarmos a discutir os modelos lineares mistos, vamos trabalhar com os dados de um experimento real, publicado em um pôster de Forster, Rodrigues & Corrêa (2019) [3]. Os dados estão na Tabela 1.1.

Partic.	Condição	Item A	Item B	Item C	Item D	Partic.	Condição	Item A	Item B	Item C	Item D
1	mult	1.17	0.64	2.81	5.69	2	sing	2.11	2.11	3.84	8.39
3	mult	2.71	2.4	2.37	6.61	4	sing	0.88	1.39	3.84	1.87
6	mult	0.76	0.8	0.84	1.49	5	sing	2.11	2.11	3.84	6.53
8	mult	2.14	0.98	0.73	6.99	7	sing	2.24	2.56	6.33	3.1
10	mult	2.01	1.63	2.59	2.88	9	sing	0.78	5.94	2.09	5.99
14	mult	1.43	1.08	1.89	3.69	11	sing	1.18	1.14	1.24	2.16
16	mult	2.79	0	0.82	2.61	13	sing	2.79	1.99	1.98	6.67
18	mult	0.41	0.72	1.42	2.54	15	sing	1.13	2.11	1.83	3.8
19	mult	1.17	1.05	1.72	3.3	17	sing	6.85	1.55	2.97	6.28
20	mult	1.02	0.88	2.34	1.47	21	sing	0.75	2.11	1.68	2.36
22	mult	2.12	1.38	2.17	5.88	23	sing	1.06	1.12	6.97	1.31
24	mult	0.94	0.66	0.78	2.01	25	sing	1.02	1.94	8.59	5.99
26	mult	1.21	1.51	2.18	0.64	27	sing	2.76	3.32	2.36	6.49
28	mult	0.84	0	0.88	0.91	29	sing	2.23	2.11	6.62	5.58
30	mult	1.08	0.71	1.41	4.22	31	sing	0.71	0.79	0.75	9.17
34	mult	1.73	2.08	5.55	5.89	33	sing	3.09	2.12	2.72	2.98
36	mult	0.75	1.9	1.56	5.92	35	sing	4.12	1.49	7.57	8.08
Médias					**2.02**						**3.29**

Tabela 1.1 – Tempos de fixação do olhar em experimento de Forster et al. (2019).

Observe que temos um *design between* sujeitos, ou seja, os sujeitos que estavam expostos à condição *mult* não estavam à condição *sing*, mas *within itens*, já que os mesmos itens - nesse caso, pranchas com desenhos de um único animal (*single*) ou mais de um animal (*multiple*) - eram vistos em ambas as condições. Repare, também, que esse é um *design* com medidas repetidas, já que, de cada sujeito eram tomadas quatro mensurações e, para cada item, duas. As médias das condições estão na última linha da tabela, mostrando que a condição *single* levava a maiores tempos para a primeira fixação ($3.29ms \times 2.02ms$). Por fim, além disso, dada o nosso debate no capítulo anterior, sabemos que temos um fator fixo (Condição, com dois níveis: *mult* e *sing*) e dois fatores aleatórios (Sujeitos e Itens).

A fim de visualizar a dispersão e organização dos dados, observe a Figura 1.1, em que são apresentados os *boxplots* para *mult* e *sing* e cada uma das observações sobrepostas em cinza.

Apenas olhando para essa imagem, podemos especular sobre algumas coisas: a condição *mult*, além de ter tempos mais baixos, parece ser bem mais homogênea do que *sing*. De fato, *sing* parece ter sua média "puxada" para cima pelos valores extremos. Observe, também, que as medianas das duas condições (a linha que divide as caixas ao meio) não se distanciam muito. Mais à frente, quando tratarmos dos diagnósticos desse modelo, essa informação será importante. Fique com ela guardada aí por enquanto.

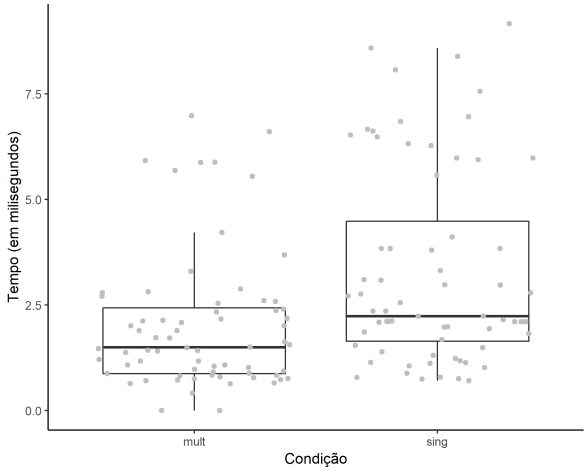

Figura 1.1 – *Boxplot* para as condições experimentais com valores observados sobrepostos

1.1 Modelos lineares: uma primeira abordagem

A fim de começarmos a analisar esses dados, vamos ajustar a eles um *modelo linear* – esse não é um modelo misto (ainda), mas um modelo simples. Nós não vamos fazer isso "na mão", como fizemos para a ANOVA – um computador vai fazê-lo pra gente; e também não vamos te explicar, agora, o que é um modelo linear. Estamos torcendo que ao longo do processo essa definição se torne mais clara.

O resultado do nosso modelo linear está na Tabela 1.2:

Resíduos				
Min	1Q	Mediana	3Q	Max
-2.5794	-1.2471	-0.6071	0.554	5.8806
Coeficientes				
	Estimativa	Erro padrão	t-valor	p-valor
Intercepto	2.02	0.243	8.315	$9.09e-14$
Condição: *sing*	1.26	0.343	3.685	0.0003

Tabela 1.2 – Modelo linear com Condição como fator fixo

Observe que temos uma tabela muito diferente da tabela da ANOVA. Não temos mais somas quadráticas nem quadrados médios nem valor de F, mas temos algumas coisas conhecidas: temos uma tabela sobre os resíduos que, obviamente, é uma tabela que descreve os quartis. Todavia, temos um negócio chamado "Intercepto" e uns coeficientes. Vamos começar por aqui então.

Observe que no Intercepto temos o valor 2.02, justamente a média da condição *mult* apresentada na tabela; e abaixo temos o valor 1.26 para Condição: *sing*. Mas essa não é a média de condição *sing*. A média para essa condição era: 3.29. Mas repare:

$$\bar{x}_{sing} - \bar{x}_{mult} = 3.29 - 2.02 = 1.26$$

O que essa conta nos mostra, portanto, é que o modelo que ajustamos está comparando as médias das duas condições, mas em vez de dizer que elas são diferentes, está mantendo uma delas fixa no intercepto (*mult*, nesse caso) e comparando a outra com ela. Com isso, os coeficientes do modelo nos dizem que a condição *sing* leva a um aumento do tempo para a primeira fixação de $1.26ms$.

Mas o que afinal é intercepto? Para entender isso, você precisa entender o que é uma equação linear. Vamos a ela. A fórmula mais básica de uma equação linear é:

$$y = ax + b$$

Desse modo, imagine que $a = 3$ e $b = 5$. Logo:

$$y = 3x + 5$$

Assim sendo, com dois valores de x, podemos montar uma reta no plano cartesiano, como mostra a Figura 1.2:

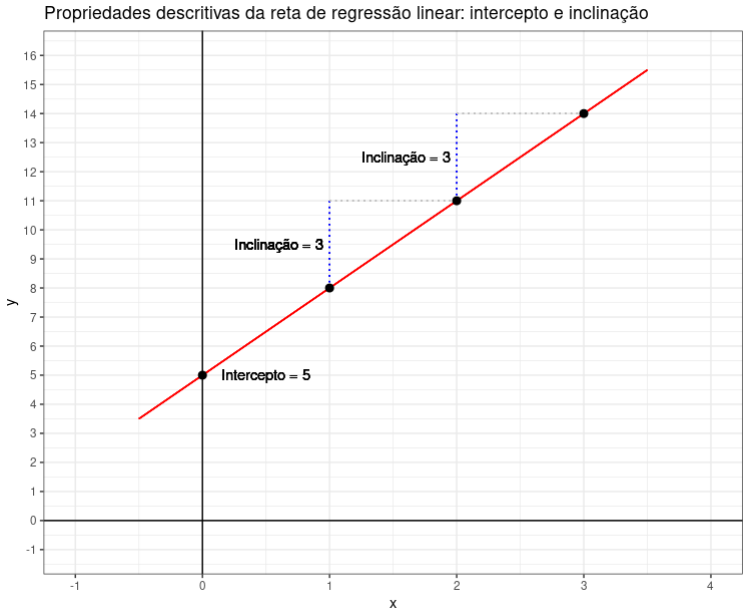

Figura 1.2 – Intercepto e inclinação na reta da equação linear

Agora observe o ponto do eixo y em que x é zero. Quando $x = 0$, $y = 5$, justamente o valor de b na equação acima. Além disso, esse ponto tem uma característica muito especial: ele é onde a reta que traçamos corta - ou intercepta - o eixo y. Ele é, por isso, chamado de intercepto.

Agora observe o ponto em que o eixo x é 1. O valor de y correspondente é 8. Vamos subir mais uma unidade no eixo x, para $x = 2$. Nesse caso, $y = 11$; para $x = 3$, $y = 14$; etc. Você reparou que sempre que subimos uma unidade em x, subimos 3 unidades em y. Olhe para a forma da nossa equação: 3 é justamente o valor que multiplica x. Esse valor é

chamado de *inclinação* (*slope*, em inglês) - o nome técnico é *coeficiente angular*, mas você não precisa saber isso.

Com isso em mente, vamos voltar para o nosso experimento e imaginar que construímos uma equação linear para ele, mas que, em vez de x, temos a nossa Condição experimental no lugar. Condição, portanto, será a nossa variável - aliás, é justamente assim que ela é chamada no *design* experimental: *variável* independente. Nossa equação terá a seguinte forma (só trocamos a e b pelas letras gregas α (alpha) e β (beta):

$$y = \alpha(\text{Condição}) + \beta$$

Nós sabemos, pela fórmula da equação linear, que β é o nosso intercepto e que α é a nossa inclinação. Nós sabemos, também, que o modelo linear colocou a condição *mult* no intercepto e calculou uma inclinação para a condição *sing*. Vamos substituí-la na fórmula, portanto:

$$y = 1.26(\text{Condição}) + 2.02$$

A nossa variável Condição pode assumir dois valores (*mult* e *sing*). A questão é que essa fórmula não adianta de nada do jeito que está, já que não podemos fazer contas com letras. Precisamos de algum modo codificar *mult* e *sing* de tal modo que possamos calcular os resultados. Isso pode ser feito porque nós sabemos, também, que o intercepto ocorre quando o valor de x é zero. E nós sabemos que o intercepto é a média de *mult*. Logo, vamos codificar a condição *mult* como zero. Se fazemos isso, obtemos:

$$y = 1.26(\text{Condição}) + 2.02$$

$$y = 1.26(\text{mult}) + 2.02$$

$$y = 1.26 \times 0 + 2.02$$

$$y = 0 + 2.02$$

$$y = 2.02$$

Observe que esse é o coeficiente estimado pelo modelo linear. Quando Condição é do tipo *mult*, então o que se espera obter em termos de tempo de fixação é $2.02ms$.

Continuando nosso raciocínio, vamos imaginar que nossa codificação é binária e que aplicamos o código 1 para condição *sing*. Nesse caso:

$$y = 1.26(\text{Condição}) + 2.02$$

$$y = 1.26(\text{sing}) + 2.02$$

$$y = 1.26 \times 1 + 2.02$$

$$y = 1.26 + 2.02$$

$$y = 3.29$$

Observe que, quando codificamos *sing* como 1, obtivemos, pelos cálculos, a média dessa condição, como você deve se lembrar da Tabela 1.1. Isso é o que o modelo está fazendo em termos algébricos, mas podemos pensar, também, em termos geométricos. Observe a Figura 1.3 gráfico abaixo:

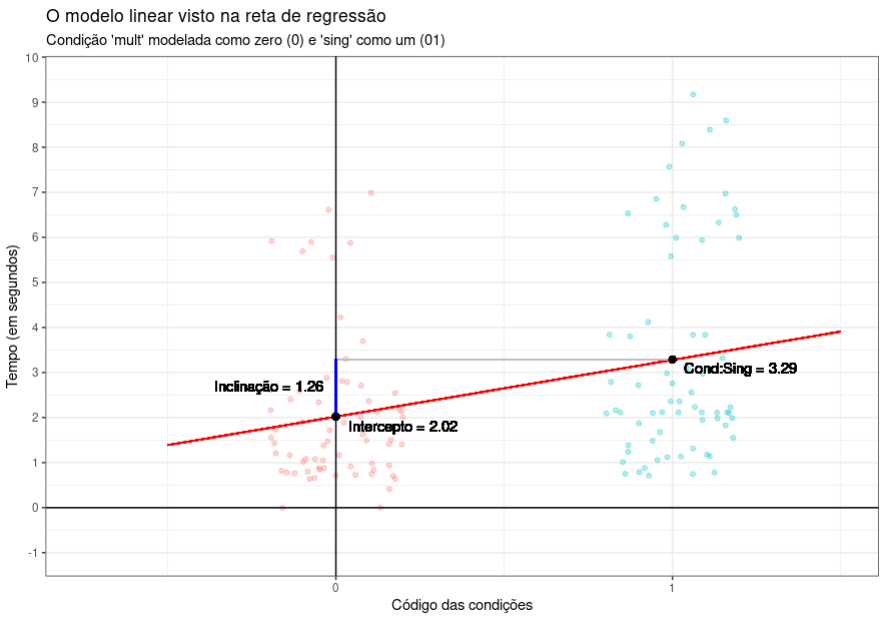

Figura 1.3 – Reta de regressão ilustrando os coeficientes do modelo linear

Perceba que a reta passa pela média da condição *mult*, ou seja, quando o eixo $x = 0$, então $y = 2.02$. Assim, quando mudamos de zero (*mult*) para 1 (*sing*), a inclinação da nossa reta, de 1.26, nos leva para 3.29, a média da condição *sing*. A esse tipo de codificação realizada pelos modelos lineares, os quais já mostramos como são algebricamente inseridos na fórmula do modelo, dá-se o nome de *constrast coding*. Não entraremos em detalhes agora sobre ele, mas futuramente você precisará saber um pouco mais sobre as diversas codificações possíveis.

Agora vamos voltar a pensarmos em termos algébricos. Um outro modo de ver e imaginar que o modelo está nos dando uma fórmula geral para prever tempos desconhecidos. Ele está nos dizendo que, se replicássemos esse experimento, é provável que o tempo dos indivíduos pudesse ser descrito por essa fórmula. Vamos, então, pegar um participante qualquer na nossa Tabela 1.1, digamos, o participante de número 10 vendo o Item B. Esse participante só viu a condição *mult* e seu tempo foi de $1.63ms$. Substituindo na fórmula:

$$1.63 = 1.26(\text{Condição}) + 2.02$$

$$1.63 = 1.26(\text{mult}) + 2.02$$

$$1.63 = 1.26 \times 0 + 2.02$$

$$1.63 = 2.02$$

Observe que nossa previsão deu errado: 1.63 não é igual a 2.02. Como já discutimos no capítulo anterior, sempre que estamos lidando com estatística, estamos lidando com variação e com erros inesperados. Na verdade, esse sujeito em particular é mais rápido $0.39ms$ do que a média das pessoas vendo essa condição: $1.63 - 2.02 = -0.39$. Por isso, precisamos incluir, no nosso modelo, um termo de erro:

$$y = 1.26(\text{Condição}) + 2.02 + Erro$$

Com isso, podemos descrever melhor o tempo mensurado para o nosso sujeito número 10 no Item B:

$$1.63 = 1.26(\text{Condição}) + 2.02 + Erro$$

$$1.63 = 1.26(\text{mult}) + 2.02 + Erro$$

$$1.63 = 1.26 \times 0 + 2.02 + Erro$$

$$1.63 = 2.02 + Erro$$

$$1.63 - 2.02 = Erro$$

$$Erro = -0.39$$

Agora imagine que você faça isso para cada um dos sujeitos vendo cada uma das condições dadas, ou seja, para cada uma das observações na Tabela 1.1 você calculou esse valor de Erro.

Isso te pareceu familiar?! Basicamente, o que estamos fazendo é calculando os desvios, como fizemos no Capítulo 1 e também como fizemos na ANOVA. Na verdade, a ANOVA é também um modelo linear, mas com uma abordagem um pouco diferente. Um resumo dos quartis desses resíduos está mostrado lá na tabela do nosso modelo. Além disso, como todo desvio, eles têm a interessante propriedade de ter média zero - já que a soma dos desvios é zero, a média também é zero, já que zero dividido por algo é zero.

Na Tabela 1.3, você pode visualizar um pouquinho do que explicamos até agora, ou seja, como cada um dos valores observados para cada sujeito, em cada condição, foi decomposto nas estimativas do modelo. Por exemplo: o valor observado de 1.17 para o sujeito 1 na condição *mult* vendo o Item A é fruto de $2.02ms$ do intercepto (um valor comum a todos), $0ms$ do fato de ser da condição *mult* e $-0.852ms$ de um fator incontrolado, o fator de Erro.

Essa tabela, no final das contas, pode ser resumida na fórmula abaixo, onde o valor observado para cada sujeito i em cada condição j é função de um valor comum a todos, o intercepto μ, um valor devido à condição j, a inclinação β_j e um fator de erro e, do sujeito i na condição j:

Partic.	Condição	Item	Tempo Observ.	Intercepto	Coeficientes	Resíduos
1	mult	A	1.17	2.022	0	-0.852
3	mult	A	2.71	2.022	0	0.688
6	mult	A	0.76	2.022	0	-1.262
			⋮			
31	sing	D	9.17	2.022	1.267	5.881
33	sing	D	2.98	2.022	1.267	-0.309
35	sing	D	8.08	2.022	1.267	4.791

Tabela 1.3 – Tempo observado em função dos valores estimados pelo modelo linear

$$y_{ij} = \mu + \beta_j X + e_{ij}$$

Isso pode parecer simples, mas tenha certeza de que você entendeu essa fórmula. Ela certamente vai facilitar muito a leitura de textos técnicos da área, que não param um segundo sequer para explicá-la para você.

Agora que você entendeu a parte dos coeficientes e dos resíduos, podemos passar ao restante dos valores mostrados na Tabela 1.2. A coluna Erro padrão simplesmente mostra o Erro padrão da nossa estimativa. Para cada coeficiente estimado, temos um erro padrão. A coluna t-valor mostra o valor da estatística de t. Do mesmo modo como na ANOVA calculamos um valor de F, uma estatística relacionada à *distribuição de Fisher-Snedecor*, nos modelos lineares teremos uma estatística relacionada à *distribuição t de Student*. Com base nesse valor, calcula-se um p-valor, que nos indica a significância da nossa estimativa.

Aqui devemos nos deter mais uma vez para explicar um detalhe. Temos dois p-valores na nossa Tabela 1.2. O primeiro está ligado ao intercepto. Mas o que ele significa? Esse p-valor está apenas nos mostrando que o coeficiente estimado para o intercepto é diferente de zero. Isso não diz muito sobre o que estamos investigando nesse experimento. Por isso, podemos ignorá-lo. No entanto, o p-valor para Condição: *sing* nos interessa: ele mostra que o coeficiente estimado para essa condição é significativo quando comparado à condição modelada no intercepto.

Há uma porção de outras coisas que poderíamos - e deveríamos - falar aqui sobre modelos lineares, mas como esse é um material didático para discutir os conceitos mais básicos, vamos encerrar por aqui e passar para os famigerados modelos mistos.

1.2 Modelos lineares mistos

Como já discutimos no capítulo sobre inferência estatística, o nosso *design experimental* tem dois fatores aleatórios e, portanto, precisamos considerá-los na nossa análise. Por isso, vamos começar essa seção ajustando um modelo linear misto para os nossos dados, incluindo nele Participantes e itens como um efeito aleatório. Daqui a pouco vamos explicar melhor a nossa estrutura de efeitos aleatórios. Os dados para o modelo ajustado estão na Tabela 1.4.

Resíduos				
Min	1Q	Mediana	3Q	Max
−2.15994	−0.57818	−0.09132	0.40515	2.99107
Efeitos aleatórios				
Grupos	Nome	Variância	Desvio padrão	
Participantes	Intercepto	0.3474	0.5894	
Itens	Intercepto	1.5878	1.2601	
Resíduos		2.4805	1.5750	
Efeitos fixos				
	Estimativa	Erro padrão	t-valor	
Intercepto	2.02	0.6737	3.001	
Condição: *sing*	1.26	0.3374	3.757	

Tabela 1.4 – Modelo linear misto com interceptos para participantes e itens

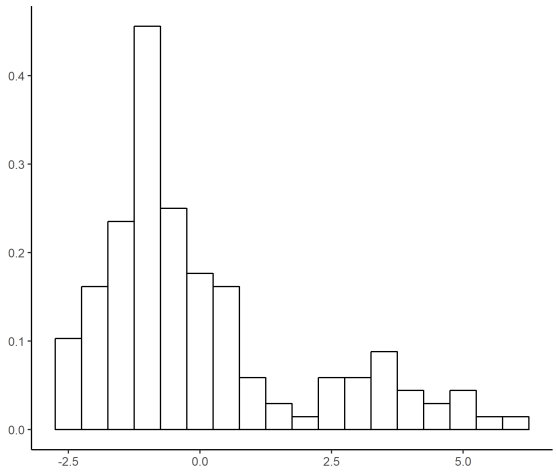

Figura 1.4 – Histograma dos resíduos do modelo simples

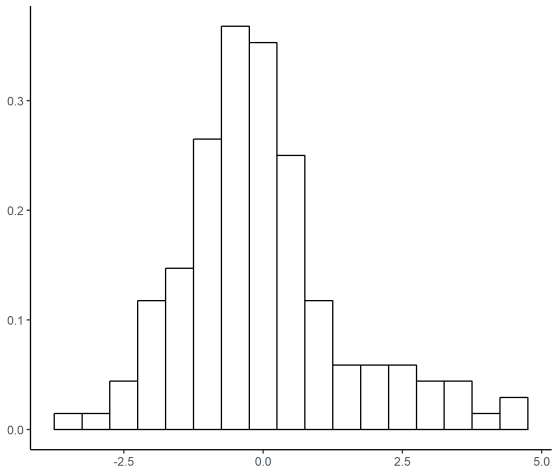

Figura 1.5 – Histograma dos resíduos do modelo misto

Observe que agora temos três conjuntos de dados: os quartis dos resíduos, os efeitos aleatórios e os efeitos fixos. Os resíduos você já sabe o que são, os efeitos fixos também. Mas repare que, apesar de os coeficientes estimados serem idênticos: 2.02 e 1.26, o Erro padrão mudou, o t-valor também e não temos mais uma coluna com o p-valor. Ademais, na parte sobre Efeitos aleatórios, temos um intercepto para participantes, um intercepto para itens e os resíduos, além da variância e do desvio padrão associado a cada um deles. Vamos tentar entender esses elementos todos.

Primeiro, já que os coeficientes estimados continuaram idênticos, comecemos investigando a estrutura dos resíduos. A Figuras 1.4 e 1.5 são de histogramas dos resíduos do modelo simples (apenas com efeito fixo) e do modelo misto (com efeito fixo e aleatórios). Como você pode perceber, a estrutura dos resíduos é bem diferente. Mas por que isso acontece? Lembre-se de que os resíduos são o fator de erro, a variabilidade não explicada

pelo modelo. O que o modelo misto está fazendo é olhar para esses resíduos e buscar estruturá-los. Nós já fizemos isso quando ajustamos ANOVAs a nossos dados. Lá, quando tínhamos um fator aleatório, calculávamos um efeito do fator aleatório. Aqui, quando temos fatores aleatórios, calculamos também o impacto desses fatores na precisão do nosso modelo. Observe com atenção a Tabela 1.5, construída nos moldes da Tabela 1.3:

Partic.	Condição	Item	Tempo Observ.	Intercepto	Coefs	Intercepto partic.	Intercepto item	Resíduos
1	mult	A	1.17	2.022	0	0.199	-0.849	-0.202
1	mult	B	0.64	2.022	0	0.199	-1.012	-0.57
1	mult	C	2.81	2.022	0	0.199	0.196	0.392
1	mult	D	5.69	2.022	0	0.199	1.665	1.804
				⋮				
33	sing	A	3.09	2.022	1.267	-0.202	-0.849	0.852
33	sing	B	2.12	2.022	1.267	-0.202	-1.012	0.044
33	sing	C	2.72	2.022	1.267	-0.202	0.196	-0.564
33	sing	D	2.98	2.022	1.267	-0.202	1.665	-1.772

Tabela 1.5 – Tempo observado em função dos valores estimados pelo modelo linear misto

Nesse caso, o valor observado para o participante 1 na condição *mult* exposto ao item A é função de um intercepto comum a todos ($2.02ms$); de um valor para a condição *mult*, que nesse caso foi modelada no intercepto, ou seja, é zero; de um valor desse participante específico, que é um pouco mais lento ($0.199ms$); de um valor devido ao fato de ser uma condição do tipo A, que é lida mais rapidamente em média ($-0.849ms$); e de um fator de erro não controlado ($-0.202ms$).

Repare que o efeito para o item A é o mesmo para o participante 33 e será o mesmo para todos os outros participantes. Esse é um efeito desse item específico. O mesmo para os demais itens.

Dizendo de outro modo, alguns participantes são mais lentos e alguns mais rápidos. Por isso, modelamos um intercepto para cada um deles, ou seja, construímos uma reta de regressão para cada um dos participantes, uma que passe horizontalmente por cada um dos pontos da Figura 1.6.

O mesmo foi feito para os itens (cada um dos 4 itens tem uma média que mostra quanto ele diverge das médias de cada condição). Se você tiver boa memória, lembrará que isso é exatamente o mesmo que fizemos nas ANOVAs: calcular o quanto de variação foi devida aos sujeitos **ou** ao itens. Aqui, porém, fizemos para os sujeitos **e** para os itens, simultanemanete. Com dissemos antes, as estimativas para os efeitos fixos continuam as mesmas, mas agora temos muito mais controle sobre o nosso fator de erro. Usando essa tabela como referência, podemos agora escrever uma nova equação para esse modelo:

$$y_{ijk} = \mu + \beta_j + \alpha_i + \gamma_k + e_{ijk}$$

Em outras palavras, o valor observado do sujeito i, na condição j, exposto ao item k é função do intercepto μ, do efeito da condição j (β_j), do efeito do participante i (α_i), do efeito do item k (γ_k) e de um fator de erro inexplicado para o sujeito i, na condição

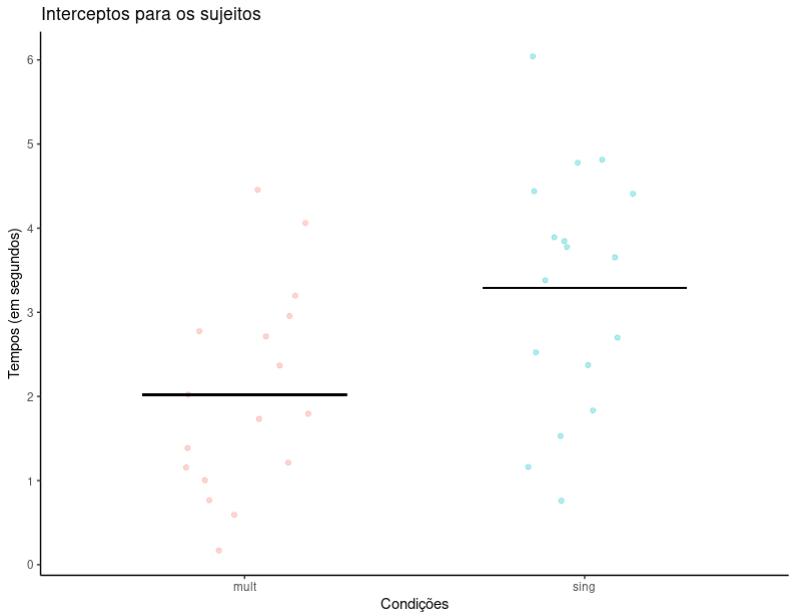

Figura 1.6 – Entendo o que fazemos quando modelamos interceptos para sujeitos

j exposto ao item k (e_{ijk}). Você pode até achar que essa informação é inútil, mas, como dissemos antes, temos confiança de que ela vai te evitar uma boa dor de cabeça quando você começar a ler os textos técnicos sobre o tema, que não têm a menor paciência para te explicar isso.

Agora que você já sabe isso, a leitura da Tabela 1.4 se torna muito mais fácil. A parte sobre efeitos aleatórios apenas mostra que calculamos interceptos para participantes, itens e o erro incontrolado (os resíduos) e suas respectivas variâncias e desvios padrão. Aqui, quanto menor a variabilidade do fator aleatório, menos impacto ele tem no nosso modelo, menos explicativo ele é dos dados. Até ao ponto em que, se a variância de um fator for muito próxima de zero, podermos excluí-lo do modelo porque, no fundo, ele não contribui em nada para a nossa análise (Mas não saia excluindo ainda. Existem regras para isso. Acalme-se!).

Observe, porém, que não consideramos ainda uma coisa. Acompanhe o raciocínio: os participantes estão aninhados (*nested*) nas condições. Logo, o sujeito que foi submetido à condição *mult* não foi submetido à condição *sing*. Logo, faz todo sentido ter um intercepto para cada um deles, mostrando o quanto eles variam em relação à média da sua condição. Os itens, no entanto, são um fator cruzado (*crossed*) com as condições. Logo, o mesmo item foi submetido tanto a *mult* quanto a *sing*. Logo, pode ser possível que um item específico gere comportamentos distintos em uma delas. Por exemplo, talvez o item A seja lido mais rapidamente em ”mult”do que o item ”B”, mas mais lentamente em ”sing”do que ”B”. Logo, eu preciso controlar não só quanto cada item varia em relação às suas médias específicas (*interceptos* para itens). Como eles são cruzados, preciso calcular o quanto eles variam por condição (*inclinações* para os itens por condição). A Figura 1.7 ilustra esse

ponto. Para os sujeitos, no entanto, não faz nenhum sentido dizer que eles se comportam de maneira distinta de acordo com a condição. Por isso, não devemos modelar *inclinações* para eles, mas apenas *interceptos*.

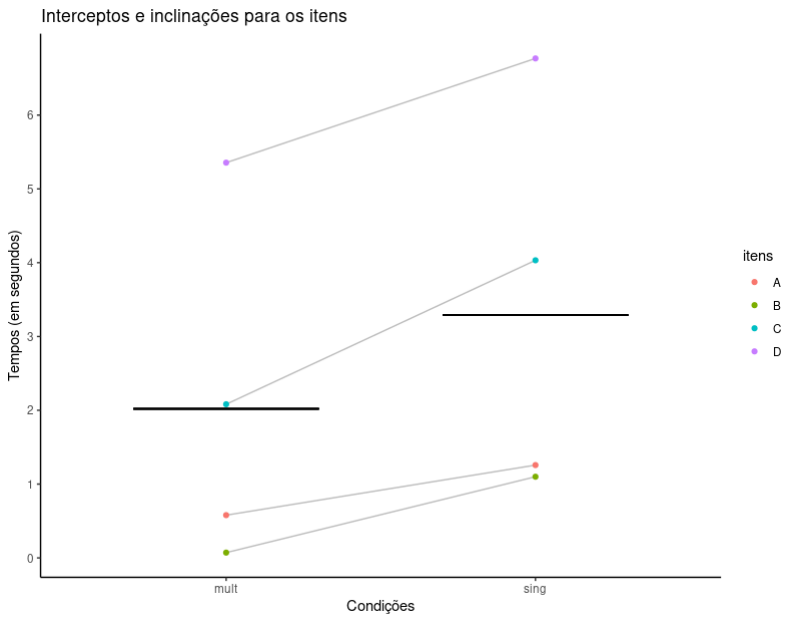

Figura 1.7 – Intercetos e inclinações (*slopes*) para os itens

Mas pela imagem podemos ver claramente que os itens são "bem comportados" com relação às condições: os itens A e C são mais rápidos do que a média e os itens B e D são mais lentos, independentemente da condição. Seja como for, nós não vamos entrar nesse modelo agora. Ele é levemente mais complexo. Vamos avançar um pouco mais para ver se conseguimos terminar esse material antes de o leitor ficar muito entediado.

Aliás, para terminar o debate, deveríamos falar sobre o p-valor, ou melhor, sobre a ausência dele. Se não temos um p-valor – e por que não temos? – como vamos saber se o nosso valor estimado é significativo? Mas vamos adiar um pouquinho esse tópico e entrar em um mais importante do que ele: nós podemos aplicar um modelo linear a nossos dados?

1.2.1 Seu modelo é saudável?

Até agora, durante todo esse material, viemos aplicando modelo atrás de modelo, olhando para os resultados e seguindo em frente. Como o material tem um propósito didático de explicar o funcionamento básico da inferência estatística, dos testes de hipótese, etc., não acreditamos que tenhamos perdido nada com isso. Mas agora chegamos a um ponto em que precisamos fazer o *diagnóstico do modelo* que ajustamos. Afinal, nós podemos aplicar esse modelo aos dados que obtivemos?

Até agora não precisamos fazer essa pergunta por que nossos dados eram artificiais, construídos de tal modo a satisfazer à maioria das exigências dos modelos até agora

aplicados. Mas agora estamos lidando com dados reais e, se quisermos fazer boa ciência, precisamos olhar com cuidado para as exigências dos modelos que estamos ajustando.

Os modelos lineares, em geral, são parte de um conjunto de análises estatísticas conhecidas como *testes paramétricos* – afinal, estamos estimando *parâmetros* populacionais a partir de *estatísticas* amostrais. Os modelos lineares, em particular, fazem algumas exigências para poderem ser aplicados com segurança. Se essas exigências não são cumpridas, não mais podemos ter confiança na qualidade da nossa estimativa – como já dissemos em outro momento, estamos fazendo adivinhação, mas adivinhação rigorosa. Nesse processo, vamos seguir o protocolo apresentado em Zuur et al. (2010) [6].

Grosso modo, os pressupostos dos modelos lineares (e também das ANOVAS, já que elas são um tipo de modelo linear) são:

1. linearidade
2. normalidade
3. homoscedasticidade ou homogeneidade de variâncias
4. independência das observações

Aqui nós precisaríamos incluir, ainda, no caso de modelos mistos complexos, o problema da *correlação*, mas vamos deixar esse caso para uma discussão à parte. Além disso, temos ainda que verificar o problema dos *outliers*, que não é um pressuposto, mas cuja presença pode enviesar nossa análise. Por fim, a questão da independência das observações foi discutida quando tratamos das ANOVAs para medidas repetidas. Portanto, não nos aprofundaremos nela por aqui.

Para o caso dos modelos lineares, a veificação dos pressupostos se dá por meio da investigação dos resíduos do modelo. Na verdade, aquela fórmula dos modelos que colocamos acima geralmente vem acompanhada do seguinte termo – $\approx N(0, \sigma^2)$ – após o termo de erro, indicando que os resíduos têm que ser aproximadamente normais, com média 0 (zero); e variância σ^2, ou seja, a variância dos resíduos tem de ser igual à variância da população.

$$y_{ijk} = \mu + \beta_j + \alpha_i + \gamma_k + e_{ijk} \approx N(0, \sigma^2)$$

A seguir, vamos discutir em mais detalhes cada um desses pressupostos e, ao longo do processo, faremos, simultaneamente, o diagnóstico do modelo que estamos ajustando.

1.2.1.1 Linearidade

Linearidade é a assunção de que os dados para os quais estamos ajustando um modelo mantêm, entre si, uma relação linear. Observe as Figuras 1.8 e 1.9. Nesse caso, apenas na primeira apresenta-se uma relação linear entre os dados. Nas demais, outros tipos de relação se apresentam.

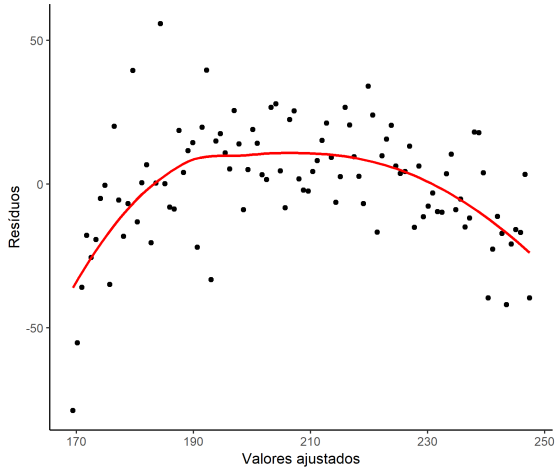

Figura 1.8 – Resíduos × ajustados mostrando relação não linear

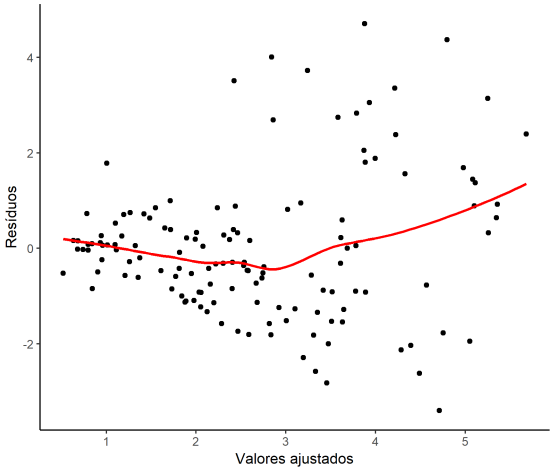

Figura 1.9 – Resíduos × ajustados para o modelo misto ajustado aos dados da Tabela 1.1

Como você pode notar, a linearidade é facilmente percebida visualmente. Para o caso dos modelos que ajustamos (mistos ou não), diz-se que o pressuposto de linearidade foi satisfeito quando os resíduos do nosso modelo são lineares. Isso é facilmente verificado quando plotamos um gráfico em que, no eixo x dispõem-se os valores ajustados pelo modelo e, no eixo y, os resíduos. Normalmente, esse gráfico é chamado, em inglês, de *residuals* × *fitted*.

Vamos discutir os dois gráficos acima. Na Figura 1.8 temos, claramente, uma relação não linear. Na verdade, a relação entre os valores ajustados e os resíduos parece ser melhor descrita por uma parábola invertida. Na Figura 1.9, por sua vez, os resíduos não apresentam uma relação parabólica e a linha vermelha começa muito próximo da média zero. No entanto, ela faz uma leve curva para baixo e depois sobe inclinadamente. Apesar de não ser de fato uma relação perfeitamente linear, parece que os dados, nesse caso, são mais lineares do que os do primeiro exemplo, já que estão mais aleatoriamente espalhados em torno da média. Vamos assumir, portanto, que a relação é próxima da linear. Observe, no entanto, que os pontos na Figura 1.9 têm uma configuração intrigante: eles apresentam um formato de cone. Daqui a pouco falaremos desse problema.

Se você quer uma regra aqui, busque uma linha reta ou aproximadamente reta e pontos aleatoriamente distribuídos em torno da média dos resíduos, como na Figura 1.10 que, obviamente, foi artificial construída. Logo, a sua não estará exatamente assim. Desse modo, se você tiver algo que caminhe para isso, terá satisfeito a condição de linearidade.

1.2.1.2 Normalidade

Normalidade é a assunção de que os dados que estamos ajustando são normalmente distribuídos, na verdade, de que a população da qual os dados foram extraídos segue

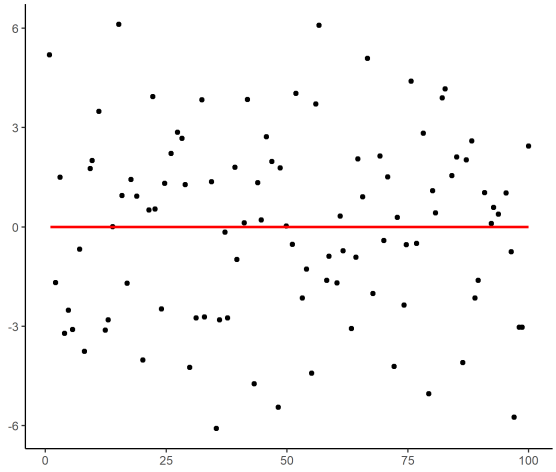

Figura 1.10 – Resíduos × ajustados mostrando o ideal a ser alcançado

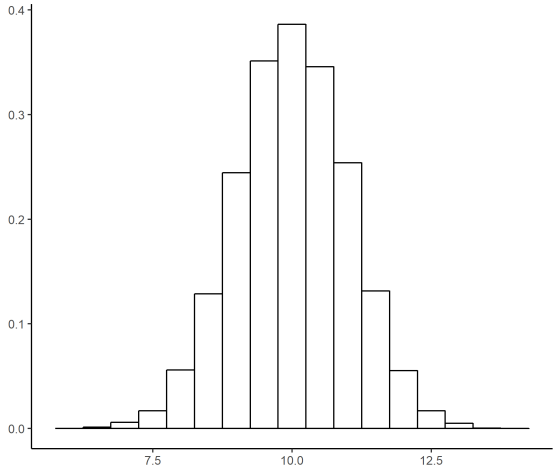

Figura 1.11 – Histograma de dados normalmente distribuídos

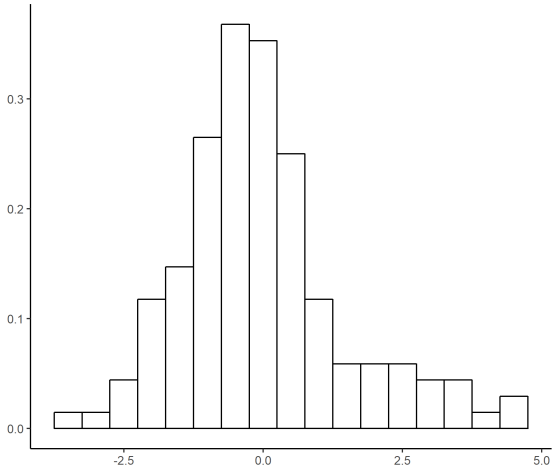

Figura 1.12 – Histograma para os resíduos do modelo misto com assimetria

uma distribuição normal. Como nos informa Zuur et al. (2010: 6-7) esse é um diagnóstico muito difícil de ser realizado na prática, mas podemos ter uma estimativa dele a partir da normalidade dos resíduos.

O método mais comum de se investigar a normalidade é plotando histogramas dos resíduos dos nossos modelos. As Figuras 1.11 e 1.12 mostram, respectivamente, um histograma de dados normalmente distribuídos e o histograma dos resíduos do modelo misto, que, como se pode ver, são assimétricos, indicando que eles não são normais.

No entanto, o método mais eficaz talvez seja plotando um gráfico em que se coloca, no eixo x os quantis de uma distribuição teórica que nós sabemos ser normalmente distribuída e, no eixo y, os quantis dos nossos resíduos. Esse tipo de gráfico é chamado de gráfico quantil-quantil ou gráfico qq (*qqplot*, em inglês). Se nossos resíduos forem

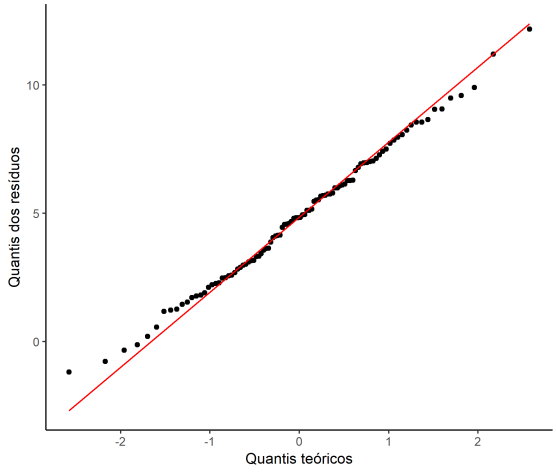

Figura 1.13 – Resíduos normalmente distribuídos

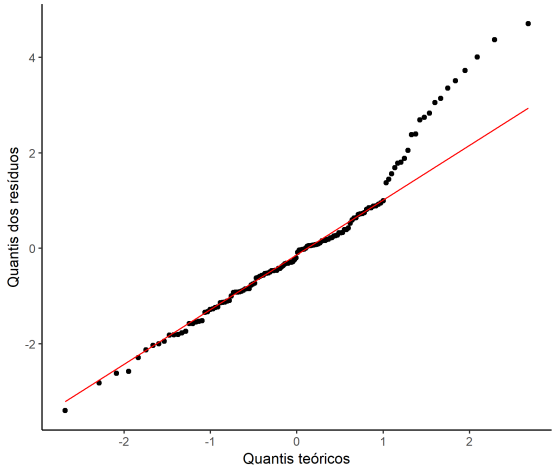

Figura 1.14 – Resíduos não normalmente distribuídos

normalmente distribuídos, os pontos se distribuirão numa linha reta, como mostra a Figura 1.13. Todavia, se nossos resíduos não forem normalmente distribuídos, isso não vai acontecer. Normalmente, o que se tem é que um – ou ambos – os extremos se afastará da linha, nos indicando que os resíduos são assimétricos naquela direção. Compare as Figuras 1.12 e 1.14, em que plotamos os resíduos do modelo misto realizado anteriormente: os pontos que se distanciam da linha vermelha nesta representam justamente a cauda mais longa à direita daquela. Logo, não há normalidade. Daqui a pouco traremos uma solução para esse problema.

Um outro caminho a se tentar quando tratamos da normalidade é realizar um *teste de normalidade*. O princípio aqui é o mesmo da realização da maioria dos testes de hipótese. Assumimos uma hipótese nula (H_0), de que os dados são normais; e apresentamos uma hipótese alternativa (H_1), de que os dados não são normais. Daí realizamos um teste. Se obtivermos um valor significativo, temos evidência para dizer que nossos dados não são normais e, logo, nosso modelo não está saudável. Existem vários desses testes prontos em pacotes computacionais, mas um dos testes mais comuns é aquele chamado de *Teste de Shapiro-Wilk* para normalidade.

Teste de Shapiro-Wilk	W	p-valor
	0.949	$7.025 \times e^{-05}$

Tabela 1.6 – Teste de Shapiro-Wilk para normalidade dos resíduos do modelo misto

Na Tabela 1.6 apresentamos o resultado desse teste para o modelo misto que ajustamos ate agora. W, nessa tabela, representa a estatística de *Wald*. Como você pode ver, temos evidência de que nossos resíduos não são normais.

Aqui precisamos deixar um adendo, no entanto. Como nos informa Zuur et al. (2010) [6], os testes de normalidade têm um grande problema: para grandes amostras, eles são

muito afetados por pequenos mudanças nos dados, de modo que se tornam pouco confiáveis; e, para pequenas amostras, o poder desses testes é pequeno, o que também os torna pouco confiáveis. A recomendação, portanto, é a de que a análise visual parece ser mais segura, de modo que você deveria confiar mais nela do que nos testes. Contudo, se o seu parecerista quiser um teste, você pode fazer um para ele.[1]

1.2.1.3 Homoscedasticidade

Homoscedasticidade é sinônimo de homogeneidade de variâncias. O seu contrário é a heteroscedasticidade. A homoscedasticidade nos diz que a variância de cada um dos fatores que estamos comparando no nosso modelo tem de ser semelhante. Assim sendo, se plotarmos um *boxplot* das nossas condições (os fatores que estamos comparando), podemos ter uma noção muito superficial quanto à variância desses fatores. Por exemplo, para o nosso experimento em questão, nós plotamos os *boxplots* apresentados na Figura 1.1, lembra-se? Se você voltar até aquela imagem, verá que a condição *mult* parece ser bem mais homogênea do que a condição *sing*. De fato, ao elevar os desvios ao quadrado para calcular a variância, nós ampliamos essa diferença. Logo, parece haver um indicativo de que os dados são heteroscedásticos.

Todavia, a análise dos *boxplots* é apenas um indicativo. A melhor maneira de investigar a homogeneidade das variância é olhando para o gráfico dos resíduos × os valores ajustados – os mesmos que usamos para detectar linearidade. Se as variâncias são homogêneas, os pontos naquele gráfico estão espalhados aleatoriamente em torno da média, não seguindo qualquer padrão detectável. Caso eles formem algum padrão, como o formato de cone da Figura 1.9, isso é sinal de que estamos diante de heteroscedasticidade. A análise daquele gráfico nos mostra claramente, portanto, que estamos diante de resíduos heteroscedásticos.

Aqui, como no caso da normalidade, você também pode realizar um teste de significância a fim de verificar se seus dados são normais. Como naquele caso, o teste é menos recomendado do que a análise visual dos gráficos, mas pode ser útil em algumas circunstâncias. Na Tabela 1.7 apresentamos o resultado do teste de Levene para homogeneidade de variâncias. Como sempre, a hipótese nula (H_0) é a de que os dados são homogêneos e a alternativa (H_1) de que não são. Pelo resultado, temos evidência para rejeitar a hipótese nula.

Um outro modo muito comum de se investigar a homogeneidade de variâncias é por

[1] A situação delineada aqui é apenas o início do problema e muito ainda precisaria ser feito para a análise dessa normalidade na prática. Como nos informa Zuur et al. (2010) *"In linear regression, we actually assume normality of all the replicate observations at a particular covariate value..."* (p.7) e *"At each covariate value, we assume that observations are normally distributed with the same spread (homogeneity). Normality and homogeneity at each covariate value cannot be verified unless many (>25) replicates per covariate value are taken, which is seldom the case in ecological studies."* (Figura 6, p.8). O ideal, portanto, seria, antes de aplicar o modelo misto, garantir que cada uma das covariantes do modelo (tratamentos, sujeitos e itens) seja normalmente distribuída e homoscedástica. Não faremos isso por enquanto.

G.L.	F	p-valor
1	9.304	0.0027
134		

Tabela 1.7 – Teste de Levene para homogeneidade de variâncias

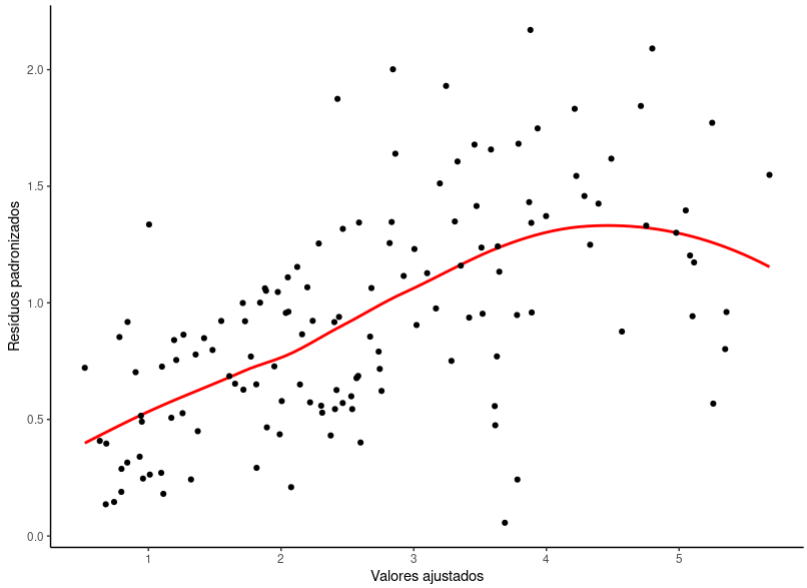

Figura 1.15 – Valores ajustados × resíduos padronizados mostrando heteroscedasticidade

meio de um *scale-location plot*. Nesse caso, em vez de plotarmos os valores ajustados (*fitted*) × os resíduos, plotamos os valores ajustados × resíduos padronizados. Resumidamente, padronizar os resíduos significa calcular a raiz quadrada dos seus módulos ou valores absolutos (o valor dos resíduos sem o sinal de + ou de − que o acompanha).

Nesse caso, estamos buscando uma linha horizontal no gráfico, que indica que as variáncias são semelhantes ao longo de toda a relação. No caso dos nossos dados, a Figura 1.15 mostra que os resíduos padronizados aumentam à medida que aumentam-se os valores ajustados. Se isso ocorre, não podemos dizer que temos homogeneidade de variâncias e nosso modelo não está saudável. Esse gráfico também será importante para identificação de *outliers*, como mostraremos mais abaixo.

1.2.1.4 Lidando com *outliers*

Até agora, descobrimos que o nosso modelo não está nada saudável e que deveria dar uma passada urgente no hospital, já que os nossos resíduos claramente não são normais e claramente não são homoscedásticos. No entanto, para terminar nosso *check-up*, precisamos ainda investigar os *outliers*, ou seja, precisamos saber se há algum valor extremo nas

observações que esteja enviesando nosso análise. Esse procedimento normalmente deve ser feito *antes* de aplicarmos o modelo e revisto *depois* que o realizamos. Contudo, como aqui estamos indo passo a passo na explicação de alguns fenômenos, vamos colocá-lo nessa etapa.

Como dissemos no capítulo introdutório, pode-se dar uma definição matemática do que seria um valor extremo (valores acima ou abaixo de um patamar, como três vezes a amplitude interquartil, por exemplo). Todavia, essa não é a melhor estratégia, sobretudo porque valores extremos podem de fato existir nos dados. Logo, precisamos pensar os *outliers* em termos teóricos.

O primeiro ponto a se pensar é que um *outlier* é um valor impróprio, ou seja, um valor advindo de um problema qualquer. Por exemplo, se um sujeito qualquer, durante a realização do experimento, começa a apertar uma tecla muito rapidamente, sem sequer ler/ver os estímulos que lhe estão sendo apresentados, é bem provável que as medidas de tempo desse sujeito devessem ser excluídas dos dados, mesmo que elas estejam dentro do limite de três vezes e meia a amplitude interquartil. O mesmo para o caso de um sujeito que se distraia e demore uma enormidade de tempo para dar sua resposta. Esse, nos parece, é ponto pacífico.

O segundo ponto a se considerar é o uso de *boxplots* para exclusão de valores extremos. Não faça isso! Não sem pensar a respeito do que você está fazendo. O que um *boxplot* mostra como *outliers* são valores que estão fora dos limites inferior e superior. Como já mostramos no capítulo sobre estatística descritiva, essa é simplesmente uma definição matemática de um valor extremo. Por exemplo, já está amplamente provado (*v.g.* Baayen & Millin, 2010) que temos de reação (e dados cronométricos em geral) não seguem uma distribuição normal, tendo uma cauda mais longa à direita. Ora, um *boxplot* de tempos de reação terá, então, alguns, senão muitos, pontos fora do limite superior, certamente aparecendo no gráficos. Entretanto, esses dados não são valores extremos, são observações reais e pertinentes que precisam ser explicadas pelo pesquisador e não excluídas. Se esse for o caso, o que seus dados precisam não é da exclusão de itens, mas de uma transformação que os deixe normalmente distribuídos – vamos falar disso na próxima seção.

Dadas essas duas ressalvas, podemos, então, investigar os valores extremos dos nossos dados. Uma boa ferramenta para começar é usar o gráfico de pontos de Cleveland [6], basicamente um gráfico de dispersão (*scatterplot*). Apresentamos quatro desses gráficos: os dois primeiros – Figuras 1.16 e 1.17 – mostram, respectivamente, uma distribuição normal sem *outliers* e uma distribuição assimétrica (*gamma*) também sem *outliers*. Observe que a distribuição assimétrica apresenta alguns valores distantes da maioria dos dados, mas isso é uma propriedade dessa distribuição, não um valor extremo real. Os demais gráficos – Figuras 1.18 e 1.19 – mostram as mesmas distribuições, agora com valores extremos adicionados artificialmente.

Observe que os valores extremos são valores que, de fato, se mostram isolados no

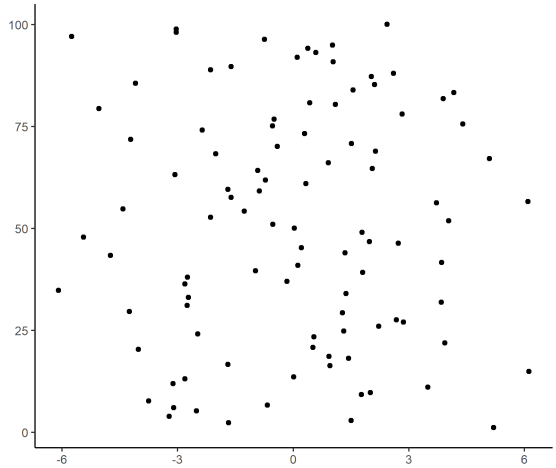

Figura 1.16 – Gráfico de dispersão para distribuição normal

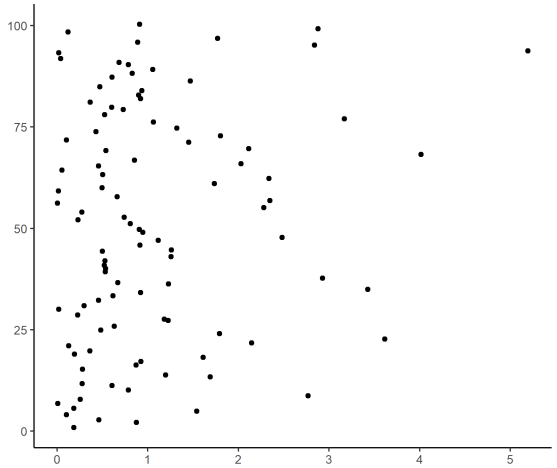

Figura 1.17 – Gráfico de dispersão para distribuição assimétrica

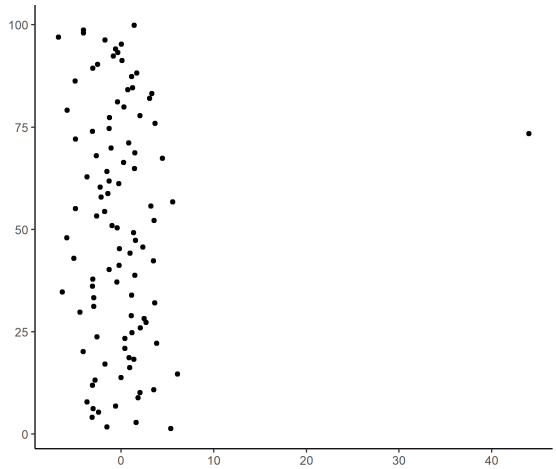

Figura 1.18 – Gráfico de dispersão para distribuição normal com possível *outlier*

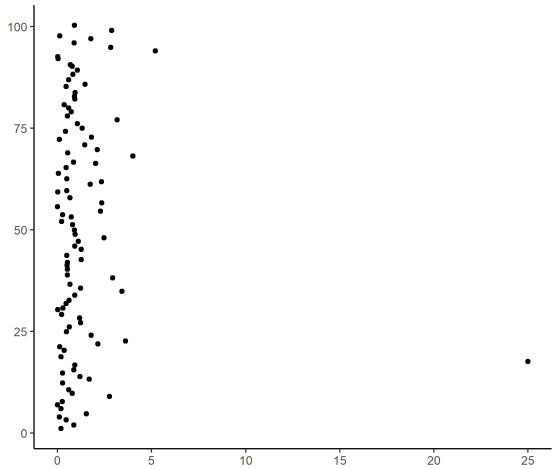

Figura 1.19 – Gráfico de dispersão para distribuição assimétrica com possível *outlier*

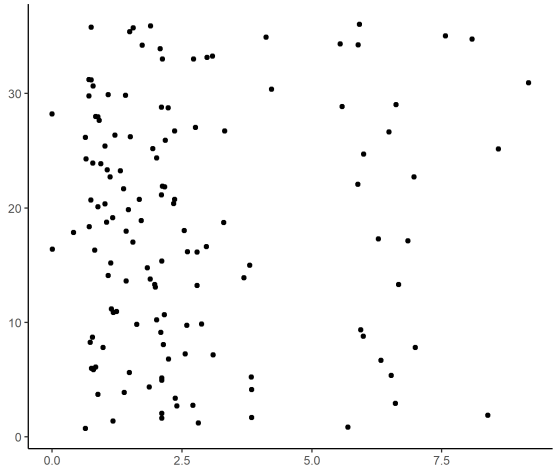

Figura 1.20 – Gráfico de dispersão para dados da Tabela 1.1

gráfico, distantes da grande massa de valores observados. Para o experimento que até aqui viemos estudando, a Figura 1.20 mostra o gráfico de Cleveland. Como você pode ver, aparentemente não temos nenhum valor extremo nos dados.

Os gráficos de dispersão que mostram os resíduos padronizados também podem ser importantes auxiliares na busca por valores extremos. Não vamos entrar em detalhes sobre as contas aqui, mas, ao padronizamos qualquer resíduo, a sua variância será sempre igual a 1. Logo, caso sejam normalmente distribuídos, algo de que precisamos nos modelos de regressão, aproximadamente 95% deles estará dentro do intervalo ± 2 (dois desvios padrão além da média) e 99% estará a até ± 3 desvios padrão da média. Logo, valores que se afastam muito desse padrão (estão muito acima de 3) são muito mais facilmente perceptíveis. Observe as figuras 1.21 e 1.22, em que são mostrados resíduos padronizados sem e com um *outlier* bem visível, para além de dois desvios padrão da média.

Uma vez que tenhamos descoberto um valor extremo, precisamos investigá-lo. Essa não é uma questão matemática, mas uma questão teórica. Pergunte-se por que aquele valor ocorreu – erro de mensuração, participante cansado, teclado com defeito, etc. – e então, tendo certeza de que aquele valor não é um valor real, exclua-o!

Um outro aspecto importante dos valores extremos é saber se eles de fato têm impacto na análise que você pretende desenvolver. Se um valor tem impacto substancial no ajuste do modelo, ou seja, se esse ponto altera as estimativas (coeficientes, desvio padrão, etc.), então se diz que esse é um *ponto de alavanca* (*leverage point*). Para os modelos lineares simples, ajustados pelo método dos mínimos quadrados, existem algumas estratégias para se fazer tal investigação, como o gráfico de resíduos $\times$ leverage e o cálculo da distância de Cook. Para os modelos mistos, no entanto, não parece haver unanimidade quanto ao melhor método a ser utilizado a fim de encontrar pontos influentes. Como, para os dados que até agora estamos investigando, não há nenhum ponto extremo, vamos considerar que

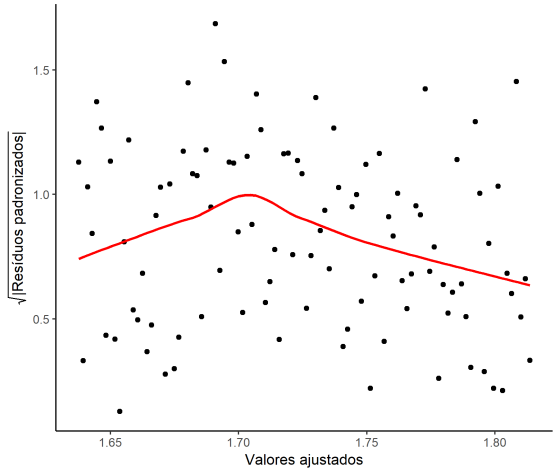

Figura 1.21 – Valores ajustados × resíduos padronizados sem *outliers*

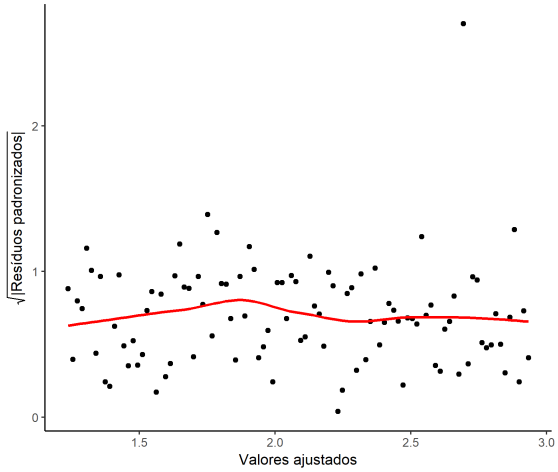

Figura 1.22 – Valores ajustados × resíduos padronizados com *outliers*

nenhum deles é influente na nossa análise e prosseguir.

1.2.1.5 A questão das transformações

Dado que chegamos até aqui, vamos fazer um pequeno resumo da análise diagnóstica do nosso modelo. Até agora verificamos que não temos valores extremos e que os resíduos são razoavelmente lineares, as boas notícias; e verificamos que nossos dados não são nem normais nem homoscedásticos, as más notícias. Desse modo, precisamos: ou mudar de modelo, para um que não leve em conta tais considerações; ou fazer alguma modificação nos dados de tal modo que eles se tornem normais e homoscedásticos. Nessa seção, vamos justamente discutir a questão das transformações.

Antes de aplicá-la, cabe um adendo: perceba que seus dados não têm um problema ou estão de algum modo errados. Eles são o que são e pensar sobre isso pode ser importante para que você entenda o fenômeno que está estudando. Por exemplo, tomemos o fato de as variâncias das condições *mult* e *sing* não serem homogêneas. Isso não é um problema, mas um achado teórico. Como observado nos *boxplots* da Figura **??**, o tempo de fixação na condição *mult* é bastante homogêneo. Todavia, quando submetidos à condição *sing*, a variação parece aumentar, o que também eleva a média. O efeito da condição, portanto, não é apenas aumentar a média, mas aumentar a variabilidade como um todo, o que pode ser importante teoricamente. Sendo isso, pense nessas questões quando estiver fazendo sua análise.

Isso tendo sido explicado, entremos um pouco na questão da transformação. Primeiro, o que é uma transformação? Uma transformação matemática é algo que qualquer pessoa realiza cotidianamente em muitas tarefas diferentes. E realizamos tão corriqueiramente que sequer lembramos que estamos fazendo uma transformação. Por exemplo, quando chegamos no mercado e pedimos 200 gramas de queijo, o rótulo que passamos no caixa

normalmente vem escrito 0.200 quilogramas de queijo e nós não achamos isso estranho. Quando fazemos uma viagem, podemos dizer que demoramos duas horas para chegar, mas também podemos dizer que viemos a uma velocidade de 100 km/hora porque estava sem trânsito. Ora, se fazemos isso, estamos simplesmente fazendo uma transformação.

Todavia, quando se trata de análise estatística, começamos a ficar angustiados, pois vemos as pessoas fazendo transformações logarítmicas, de raiz quadrada, hiperbólicas de primeiro grau, gamma, inversa da normal e uma porção de outros nomes que nos parece indicar que existe algo esquisito acontecendo por aí. As pessoas ajustam um modelo, descobrem que os dados não são normais e então aplicam uma operação de nome matemático esquisito sobre essas dados. Desculpe-me, mas isso me cheira a trapaça. Você não obteve os resultados que queria e agora está manipulando seus dados de tal modo a conseguir um p-valor e ser publicado, seu tratante!

Na verdade, não é nada disso. Isso porque um aspecto importante das transformações é que elas não alteram a natureza do que está sendo medido. Transformações são procedimentos matemáticos, não procedimentos sobre o mundo material. Por exemplo, quando alguém nos pergunta a distância entre Rio e São Paulo, normalmente dizemos que é de algumas centenas de quilômetros; mas, se nos perguntam a distância até o hospital na esquina, dizemos que é de alguns metros. Poderíamos, no entanto, informar a distância até São Paulo em metros e a distância até o hospital em quilômetros – para desespero do nosso interlocutor, obviamente. Apesar do absurdo, em ambos os casos estamos dando a mesma informação. A distância não muda devido à unidade que usamos. Tanto não muda que posso dizer que alguém tem mil reais na conta ou cem mil centavos. Isso significa que ela tem mais dinheiro no segundo caso? Obviamente não. Dizer que o Rio de Janeiro no verão faz 40° Celsius ou 104° Fahrenheit ou 313.15° Kelvin não muda a temperatura em um grau sequer. Nós estamos acostumados com essas transformações e elas não nos parecem estranhas.

Pense sobre isso: eu posso dizer que o tempo de reação a uma palavra qualquer é de 200 milisegundos, ou que é de 5.29 logaritmo natural de milisegundos, ou que é igual a $2.718282^{5.298317}$ milisegundos, ou que é igual a $\frac{2000}{10}$ milisegundos, etc. Isso não muda a minha medida. Isso muda o modo como eu expresso a minha medida. Mas, se não muda a minha medida, por que, afinal, aplicar uma transformação a meus dados? Ora, porque muda a distribuição dos seus dados. Vamos verificar isso usando a transformação logarítmica.

1.2.1.5.1 A transformação logarítmica

Como você deve ser lembar do Ensino Médio – se não ficou muito traumatizado a ponto de recalcar freudianamente essa informação –, a função logarítmica simplesmente nos dá certos expoentes (o logaritmo) a que um número (a base) tem de ser elevada para obtermos o logaritmando . Por exemplo, se o logaritmo de a, na base b, é igual a c

($log_b a = c$), então, $b^c = a$, sendo c o logaritmo, ou seja, o expoente que procuramos. Como você deve se lembrar, existem três tipos de bases muito comuns: o logaritmo de base 10 (log_{10}), o de base 2 (log_2) e o logaritmo natural, de base e, sendo $e = 2,7182818284...$, o chamado *número de Euler*.

Para o que nos interessa aqui, vamos começar observando o conjunto de dados v e o logaritmo na base 10 de cada um dos valores de v, que chamamos de $log_{10}(v)$. Como dissemos acima, os valores apresentados no segundo conjunto são simplesmente os expoentes a que elevamos 10, a base, para obtermos cada um dos números do primeiro conjunto (v). Assim: $10^0 = 1$, $10^1 = 10$, $10^2 = 100$, etc.

$$v = \{1, 10, 100, 1000, 10000\}$$

$$log_{10}(v) = \{0, 1, 2, 3, 4\}$$

Pense um pouquinho sobre o que acabamos de fazer em termos geométricos: a distância entre cada um dos números de v era sempre multiplicada por dez, ou seja, a distância entre o último e o penúltimo ponto era de 9000 unidades ($10000 - 1000$); e a distância entre o penúltimo e o antepenúltimo era de 900 unidades ($1000 - 100$). O que o logaritmo fez? Reduziu essa distância para uma (01) unidade ($4 - 3 = 1$ e $3 - 2 = 1$). Pensando metaforicamente, poderíamos dizer que ele "encolheu" as distâncias, "achatando" a longa cauda à direita dos valores originais.

Agora pense nisso em termos das distribuições de frequência. Imagine que temos uma distribuição assimétrica, com uma longa cauda para os valores positivos, como a distribuição dos tempos de reação no nosso experimento, por exemplo. Se aplicamos uma transformação logarítmica a ela, veremos que ela vai "comprimir" os valores maiores, trazendo-os mais para perto da maioria dos dados, de modo que conseguimos transformar uma distribuição assimétrica em uma distribuição simétrica. Por isso se diz que a transformação logarítmica tende a *normalizar os dados*.

Veja as Figuras 1.23 e 1.24, que mostram exatamente isso: os histogramas e curvas de densidade de uma distribuição assimétrica (os tempos em segundos do nosso experimento, que claramente têm uma cauda longa à direita)e os tempos das mesmas distribuições após uma transformação logarítmica (nesse caso, usando o logaritmo natural, o de base e).

Agora imagine que temos um conjunto de dados que foi dividido em dois – suponhamos, duas condições experimentais – e que um desses conjuntos tenha uma maior variância do que o outro (basicamente, o problema da heteroscedasticidade, já discutido). Ter uma variância maior significa, basicamente, que os valores de um conjunto são mais espalhados, dispersos, do que os valores de outro. Assim, quando aplicamos a função logarítmica a esse conjunto de dados, ela vai afetar mais os valores extremos do conjunto mais disperso, "comprimindo-os" mais, para usar a metáfora de antes. Ao fazer isso, ela tenderá a deixar

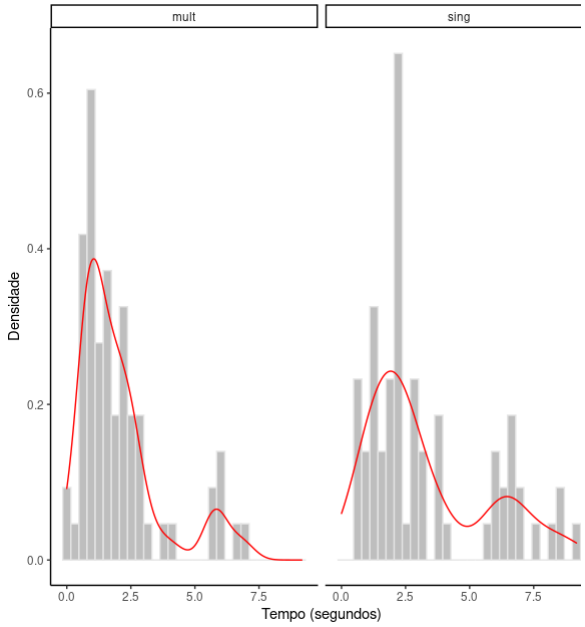

Figura 1.23 – Distribuições assimétricas com caudas à direita

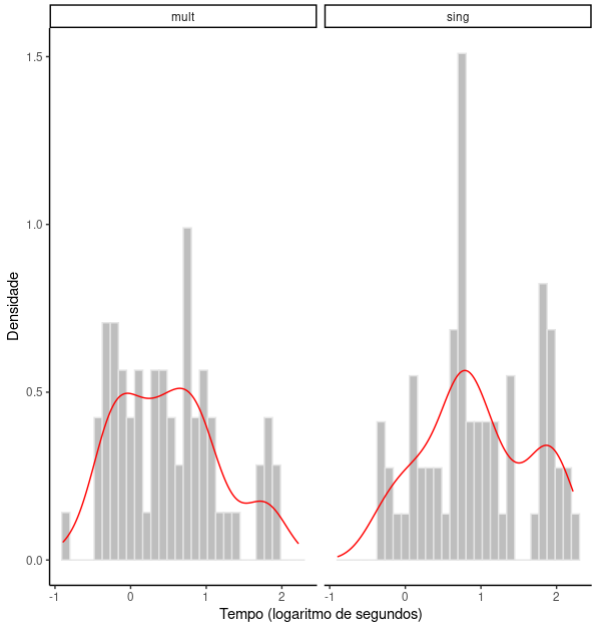

Figura 1.24 – Distribuições mais simétricas com caudas ”comprimidas”pela transformação logaritma

a dispersão dos dados parecida. Por isso, diz-se, também, que a transformação logarítmica *estabiliza as variâncias*.

Já que chegamos até aqui, até ao ponto em percebemos o quanto a transformação logarítmica é importante na psicolinguística – ver, por exemplo, Baayen & Milin (2010) [] –, e que vamos precisar desse conhecimento algumas páginas abaixo, vamos continuar para um interessante propriedade dos logaritmos que será muito útil para o caso de tempos obtidos com rastreamento ocular. Observe o conjunto de dados v' $(v - linha)$ e os seus logaritmos na base 10:

$$v' = \{0.0001, 0.001, 0.01, 0.1\}$$

$$log_{10}v' = \{-4, -3, -2, -1\}$$

O que desejamos mostrar com o exemplo acima é que, quando aplicamos a função logaritmo a um número muito pequeno, no intervalo entre 0 e 1, ela funciona, de certa maneira, de modo invertido. Vamos pensar em termos de plano cartesiano. Para valores maiores ou iguais a 1, o logaritmo transformou os dados e os distribuiu simetricamente no eixo dos números inteiros positivos $(1, 2, 3, 4,$ etc.); para os valores entre 0 e 1, o logaritmo transformou os dados e os distribuiu simetricamente no eixo dos inteiros negativos $(-1, -2, -3, -4,$ etc.). Metaforicamente, poderíamos dizer que a função logarítmica está igualando o tamanho de dois infinitos: o infinito de 0 a 1 agora tem o mesmo “tamanho” que o infinito de 1 ao... infinito.

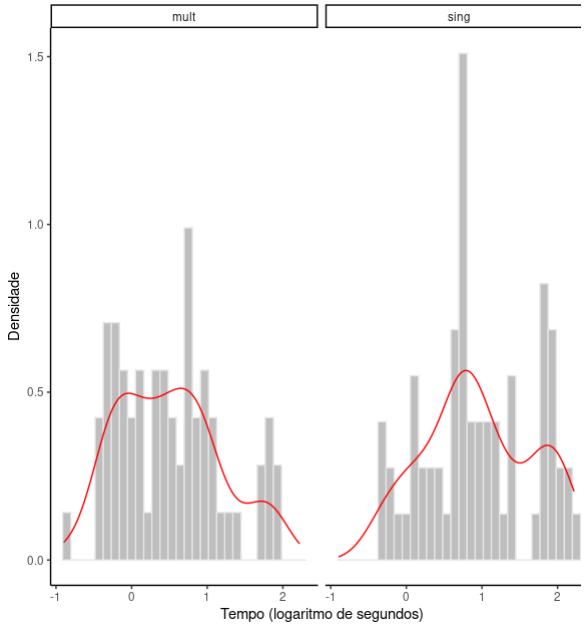

Figura 1.25 – Observe no eixo X os valores negativos

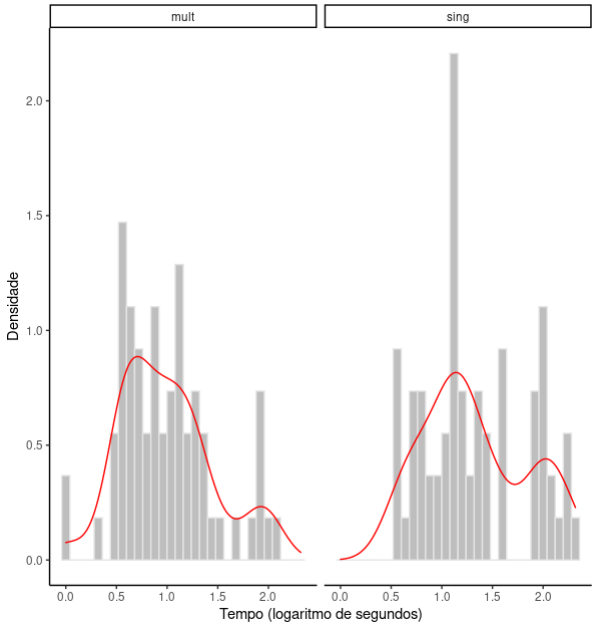

Figura 1.26 – Aqui o eixo X começa em zero

Mas por que isso nos interessa? Como informa Osborne (2008)[4], quando se trata de transformações com números pequenos, é preciso investigar esses efeitos. Para o caso em questão, se todo nossos dados são maiores do que 1, então a transformação logarítmica está fazendo a mesma coisa com todos eles, "comprimindo-os" na mesma direção. Contudo, se temos dados que estão tanto no intervalo entre 0 e 1 quando acima de 1, então a função logarítmica vai fazer coisas diferentes com cada um desses números, de modo que pode eliminar, de fato, as diferenças que estamos procurando. E isso é exatamente o caso do experimento que estamos analisando aqui. O tempo para a primeira fixação do olhar é muito pequeno, alguns da ordem de meio segundo ($0.5seg$, ou seja, entre 0 e 1) e alguns um pouco maiores, da ordem de $2, 3, 5seg$ (acima de 1). Observe as figuras abaixo: a Figura 1.25 contém os dados em logaritmos de segundos; e a Figura 1.26 contém os mesmos dados transformados (entraremos em detalhes quanto a essa transformação na próxima seção):

Um último comentário quanto à transformação logarítmica que pode te dar nos nervos quando estiver realizando uma análise com um conjunto grande de dados e de repente perceber que seu computador não executa certas ações. Como vimos acima, quando mais próximo de zero um valor, menor (mais próximo do infinito negativo) é o logaritmo. Contudo, o logaritmo de zero não é definido matematicamente, afinal, qual o expoente que elevado a qualquer base nos dá zero? Dadas as propriedades da potenciação, esse número não existe. Assim sendo, se você tiver um zero na sua tabela de dados e precisar realizar a transformação logarítmica, então esses valores precisarão ser excluídos. Coincidentemente, é justamente o caso do nosso experimento. Então, vamos a ele.

1.2.2 Valores perdidos ou *missing values*

Quando mostramos a Tabela 1.1, ela estava completa. A verdade, no entanto, é que, quando se trabalha com experimentos reais, às vezes pode ocorrer de uma medição se perder. Se você voltar àquela tabela, verá que os participantes 16 e 28 tiveram um tempo de fixação igual a zero, o que é irrealístico. Provavelmente esse foi um dado perdido pelo computador. Mas a situação é bem pior do que isso. Fomos informados pelos pesquisadores que vários outros valores daquela tabela foram perdidos. Desse modo, decidimos, para fins didáticos, substituí-los, lá naquela tabela, pela média de cada item. Abaixo, então, vamos reapresentá-la, agora com as lacunas para os valores perdidos (prepare-se, pois é de doer o coração).

Partic.	Condição	Item A	Item B	Item C	Item D	Partic.	Condição	Item A	Item B	Item C	Item D
1	mult	1.17	0.64	2.81	5.69	2	sing				8.39
3	mult	2.71	2.4	2.37	6.61	4	sing	0.88	1.39		1.87
6	mult	0.76	0.8	0.84	1.49	5	sing				6.53
8	mult	2.14	0.98	0.73	6.99	7	sing	2.24	2.56	6.33	3.1
10	mult	2.01	1.63	2.59	2.88	9	sing	0.78	5.94	2.09	5.99
14	mult					11	sing	1.18	1.14	1.24	2.16
16	mult	2.79	0	0.82	2.61	13	sing	2.79	1.99	1.98	6.67
18	mult	0.41	0.72	1.42	2.54	15	sing	1.13		1.83	3.8
19	mult	1.17	1.05	1.72	3.3	17	sing	6.85	1.55	2.97	6.28
20	mult	1.02	0.88	2.34	1.47	21	sing	0.75		1.68	2.36
22	mult	2.12	1.38	2.17	5.88	23	sing	1.06	1.12	6.97	1.31
24	mult	0.94	0.66	0.78	2.01	25	sing	1.02	1.94	8.59	5.99
26	mult	1.21	1.51	2.18	0.64	27	sing	2.76	3.32	2.36	6.49
28	mult	0.84	0	0.88	0.91	29	sing	2.23		6.62	5.58
30	mult	1.08	0.71	1.41	4.22	31	sing	0.71	0.79	0.75	9.17
34	mult	1.73	2.08	5.55	5.89	33	sing	3.09	2.12	2.72	2.98
36	mult	0.75	1.9	1.56	5.92	35	sing	4.12	1.49	7.57	8.08
Médias					**2.08**						**3.40**

Tabela 1.8 – Tempos de fixação do olhar em experimento de Forster et al. (2019).

Como você pode ver, perdemos todos os dados do participante 14 (4 observações), três dos participantes 2 e 5 (6 observações), e 1 dos participantes 15, 21 e 29 (3 observações). Contando com os 2 zeros dos participantes 16 e 28 – que não tiramos apenas por consolo – , temos uma perda de 15 observações do total de 136, ou 11.02% dos dados, ou seja, uma perda drástica. As médias também mudaram levemente.

Aqui, se estivéssemos trabalhando com ANOVA, já poderíamos estar chorando de cócoras no chuveiro. Isso porque, em uma ANOVA com medidas repetidas, não poderíamos ter lacunas na medição de um sujeito. Se perdemos uma medição, teríamos que retirar todas as observações daquele sujeito. Uma estratégia que se usava para resolver esse problema da perda de dados era, para cada participante, calcular a sua média para todas as medidas repetidas desse participante e então ajustar a ANOVA com base nessas médias. Com isso, o que se fazia era transformar uma ANOVA de medidas repetidas em uma ANOVA para amostras independentes. Observe que isso, porém, promove uma redução drástica da variabilidade dos sujeitos e dos itens. Como vimos anteriormente, com os modelos mistos, podemos considerar justamente a variabilidade desses fatores aleatórios e,

com isso, controlar melhor o nosso termo de erro.

Essa é uma outra grande vantagem dos modelos mistos sobre a ANOVA, já que os modelos mistos não têm qualquer problema em trabalhar com dados perdidos. No caso da Tabela 1.8, podemos ajustar um modelo misto para ela sem problemas.

No entanto, como vimos que a transformação logarítmica resolvia o problema da normalidade e da heteroscedasticidade, vamos aplicar o modelo nos dados transformados. O resultado está compilado na tabela 1.9.

Resíduos				
Min	1Q	Mediana	3Q	Max
−2.53666	−0.59446	−0.06351	0.60624	2.15946
Efeitos aleatórios				
Grupos	Nome	Variância	Desvio padrão	
Participantes	Intercepto	0.1055	0.3248	
Itens	Intercepto	0.1862	0.4315	
Resíduos		0.2793	0.5285	
Efeitos fixos				
	Estimativa	Erro padrão	t-valor	
Intercepto	0.4720	0.2402	1.966	
Condição: *sing*	0.4485	0.1500	2.989	

Tabela 1.9 – Modelo linear misto com interceptos para participantes e itens – estimativas em logaritmo natural de milisegundos

Perceba que as nossas estimativas agora estão apresentadas em logaritmo de milisegundos e não mais em milisegundos, o que dificulta em muito a interpretação. No entanto, perceba que os valores são muito próximos: 0.47 e 0.44 logaritmos de milisegundos. Mas o quanto isso é próximo em termos de milisegundos de verdade? Se quisermos converter os logaritmos de volta em uma escala interpretável, podemos aplicar a função exponencial às nossas estimativas. Para esses dados, exp(0.4720) e exp(0.4485) são, respectivamente, 1.60 e 1.56. Perceba que esse valor não é representativo das médias das nossas condições e nem das estimativas que tínhamos antes. Na verdade, como o intercepto mostra a condição *mult* e Cond:*sing*, a diferença entre as condições, então *sing* deveria ser bem menor do que *mult*, mas não é.

Se você voltar ao final da última seção, entenderá por que isso ocorreu. Nós aplicamos uma função logarítmica a dados que não devíamos. Portanto, vamos fazer as coisas direito e, antes de fazer a transformação, somar 1 a todos os valores – a recomendação de Osborne (2008) [4] para esses casos[2]. Se fizermos isso, o menor valor (0.41), do sujeito 18 exposto

[2] Lembre-se, esse é um material didático que busca lhe mostrar alguns problemas possíveis. No caso dos dados aqui, talvez a melhor solução fosse aplicar um outro tipo de transformação: a conversão de segundos para milisegundos. Isso certamente resolveria o problema dos valores entre zero e um. Que outras questões surgiriam, deixamos para você investigar.

ao item A, se tornará 1.41 e poderemos aplicar a transformação normalmente. Abaixo está a tabela desse modelo:

Resíduos				
Min	1Q	Mediana	3Q	Max
−2.2290	−0.6880	−0.1062	0.5697	2.3006
Efeitos aleatórios				
Grupos	Nome	Variância	Desvio padrão	
Participantes	Intercepto	0.04348	0.2085	
Itens	Intercepto	0.08954	0.2992	
Resíduos		0.13351	0.3654	
Efeitos fixos				
	Estimativa	Erro padrão	t-valor	
Intercepto	1.00789	0.16514	6.103	
Condição: *sing*	0.30480	0.09949	3.063	

Tabela 1.10 – Modelo linear misto com interceptos para participantes e itens – estimativas em logaritmo natural de milisegundos com tempo acrescido de 1

Agora, só de olhar para as estimativas, já ficamos até mais aliviados. Quando aplicamos a função exponencial àquelas estimativas, obtemos: exp(1.00789)=2.74 e exp(0.30480)=1.35, que nos informa que o intercepto (a condição *mult*) teve um tempo médio estimado pelo modelo de $2.74ms$ e que a condição *sing* tende a aumentar esse tempo de $1.35ms$. Logo, o tempo médio dessa condição é de $2.74 + 1.35 = 4.09$. Observe que esses valores não são idênticos – na verdade são bem distintos – dos valores apresentados na Tabela 1.8. Isso ocorreu porque o modelo está "estimando" o tempo da população a partir dos dados da amostra, não simplesmente "copiando" o que a amostra nos dá. Como ele faz isso: através de um método de estimação chamado de Máxima Verossimilhança Restrita (*Restricted Maximum Likelihood* ou REML), mas isso é tópico para um outra conversa.

Já que decidimos ficar com esse modelo, precisamos verificar a sua saúde. Nesse ponto, você terá que começar tudo de novo, avaliando normalidade, homoscedasticidade, linearidade e independência das observações. Como nosso propósito aqui é apenas didático, não vamos fazer tudo de novo, mas você deveria, pelo menos para treinar um pouco...

1.2.3 Intervalo de confiança, t-valor, p-valor e razão de verossimilhança

Vamos assumir, por ora, que esse é o nosso modelo final e que será esse que iremos publicar. Agora que chegamos até aqui, como reportar nossos dados?

O primeiro problema é que até agora não falamos nada sobre o famigerado p-valor. E isso foi proposital, já que nenhum estatístico parece saber, até o momento, como calcular um p-valor para os modelos mistos. Na verdade, existe uma série de propostas, mas nenhum consenso, de modo que, enquanto os estatísticos brigam sobre o tema, nós temos que continuar mesmo sem o p-valor.

Isso, porém, não é de todo ruim, já que nos sustentarmos em apenas um número e procuramos por ele como um Graal mágico e salvador não parece ser muito saudável. Aliás, existe todo um debate entre os especialistas sobre a questão do p-valor e o que de fato ele representa e se devia ou não continuar sendo usado. Uma boa leitura para você se introduzir no tema é Winter (2019) [5]. Sendo assim, vamos dar três opções para você reportar sua análise: o uso da estatística t fornecida pelo modelo, o uso de intervalos de confiança para as estimativas – provavelmente a mais recomendada – e também – com algumas ressalvas – o uso de p-valor.

Para começar, recomendamos que você leia a discussão em Baayen, Davidson & Bates (2008: 396+)[1] sobre por que é tão complicado encontrar um p-valor "correto" para os modelos mistos. Na verdade, há uma série de possíveis modos de calcular o p-valor dos modelos mistos, mas nenhum deles parece ser o "correto". Isso porque não há um modo preciso de estimar os parâmetros populacionais que estão sendo incorporados em um modelo com complexa estrutura de efeitos aleatórios. Com isso, se não sabemos sequer quantos parâmetros estamos estimando, como saber quantos graus de liberdade temos num modelo? Mesmo que isso fosse possível, para modelos desbalanceados, com muitos valores perdidos – como o que acabamos de ajustar – , não se pode sequer associar uma distribuição, como a t de Student ou a F de Fisher-Snedecor aos dados. Desse modo, torna-se inviável calcular um p-valor que seja preciso.

Esse é o problema. Mas qual a solução? A primeira é considerar a estatística t apresentada no sumário do modelo. Apesar de não sabermos calcular com precisão os graus de liberdade de um modelo misto, em geral esses valores são altos. Logo, podemos considerar a estatística t com altos graus de liberdade e estimar a significância do modelo a partir disso (ver nota 1 em Baayen, Davidson & Bates (2008)[1]). Como regra geral, se t for maior do que 1.96, então temos confiança, com $\alpha = 0.05$, de que nossa estimativa é significativa. Pedindo para meu computador calcular aqui – mas você pode procurar numa tabela da distribuição t também – , com 1 grau de liberdade no numerador e 250 no denominador, o valor de t e os respectivos índices de significância são:

t valor	p valor
1.960	0.050
2.054	0.040
2.326	0.020
2.576	0.010
3.090	0.002
3.291	0.001

Tabela 1.11 – Valor da estatística de t com 250 graus de liberdade e p valores correspondentes

Olhando para a estatística t do nosso modelo, podemos ver que $t = 3.063$, ou seja, que o coeficiente estimado é significativo com $\alpha = 0.01$. Logo, poderíamos reportar esses dados

da seguinte maneira:

Ajustou-se aos dados um modelo linear misto, com condição como fator fixo, intercepto para sujeitos e itens como fatores aleatórios, de modo que o tratamento sing *mostrou-se significativo* (sing: 0.30480 log(seg); t(1, 250)=3.063; p=0.01)*, indicando que essa condição tende a aumentar o tempo para a primeira fixação em relação à condição* mult*.*

Mas o leitor pode estar muito incomodado com essa sugestão, já que estamos "chutando" um valor de t sem qualquer critério e isso não parece muito rigoroso matematicamente. Então vamos tentar uma técnica mais precisa.

Um segundo modo de reportar esses dados seria calcular um intervalo de confiança para as estimativas do modelo. Esse talvez seja o modo mais comum de reportar as estimativas de um modelo misto, visto que o próprio modelo já nos fornece valores para o Erro padrão das estimativas. Como o intervalo de confiança é baseado nesse valor, então bastaria estabelecermos um nível de significância (α). Os pacotes computacionais que ajustam modelos mistos geralmente têm modos de realizar isso rapidamente. Para o caso em questão, nosso intervalo de confiança para o coeficiente estimado da condição *sing* é: $0.108, 0.499$.

Ajustou-se aos dados um modelo linear misto, com condição como fator fixo, intercepto para sujeitos e itens como fatores aleatórios, de modo que o valor estimado para a condição sing *foi de* 0.30480 *log(seg) com intervalo de confiança (com nível de significância 0.05) de* $0.108, 0.499$*, indicando que essa condição tende a aumentar o tempo para a primeira fixação em relação à condição* mult*.*

Se você desejar, pode apresentar um gráfico dessa estimativa, como na Figura 1.27.

Por fim, se você quiser – ou precisar muito, devido a algum parecerista mais insistente – você pode dar um jeito de arrumar um p-valor para o seu modelo. Os pacotes computacionais também têm métodos para isso, mas, como dissemos antes, nenhum deles é unanimemente aceito pelos estatísticos. Usando o pacote "afex", no R, para o modelo em questão, obtive um p-valor igual a 0.00440 Com isso, podemos reportar como fizemos antes, com o uso da estatística t.

Por fim, você pode fazer um teste de comparação de modelos. Esse é talvez o método mais comum nas publicações a respeito. Resumidamente, ele consiste em criar um modelo sem o fator que você está testando (*modelo nulo*) e um modelo com esse fator (*modelo cheio*) e então submeter esses modelos a um teste de razão de verossimilhança. A lógica por trás desse pensamento é a seguinte: ao ajustar o modelo nulo, você estará ajustando um modelo matemático sem considerar os tratamentos; e ao ajustar o modelo cheio, você estará

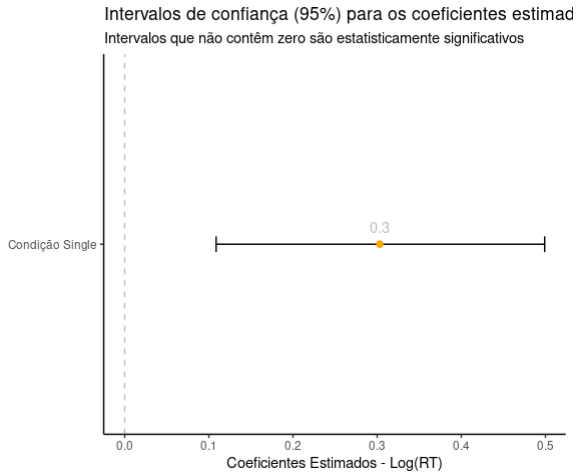

Figura 1.27 – Apresentando um gráfico da sua estimativa

considerando esses tratamentos. Quando comparar os dois modelos, você estará, então, verificando qual deles melhor se adequa aos dados. Se os tratamentos foram importantes para a explicação dos dados, então o modelo cheio apresentará uma redução em certos valores e também um valor significativo. Observe a Tabela 1.12, em que se apresenta uma comparação desse tipo.

	npar	AIC	BIC	logLik	deviance	Chisq	G.L.	p valor
Modelo nulo	4	150.42	161.61	-71.211	142.42			
Modelo cheio	5	143.77	157.75	-66.886	133.77	8.6502	1	0.00327

Tabela 1.12 – Teste de razão de verossimilhança (*likelihood ratio test*)

Não vamos entrar em detalhes sobre esse teste. Mas observe que a hipótese nula nos informa que os modelos são igualmente bons em descrever os dados. Logo, temos evidência ($p = 0.03$) em favor de negar a hipótese nula. Além disso, algumas métricas apresentadas aí nos mostram que o modelo cheio é melhor, já que seu BIC (*Bayesian Information Criterion*), seu AIC (*Akaike Information Criterion*) e seu desvio (*deviance*) são menores do que o do modelo nulo. Logo, esse p-valor poderia ser reportado do mesmo modo como foram reportados os p valores já tratados anteriormente.

Aqui, se você quiser, e puder, e achar necessário, acrescente também um gráfico de barras ou semelhante com as médias. Para o caso em questão, achamos desnecessário, já que temos apenas dois tratamentos. Em modelos mais complexos, porém, torna-se inevitável acrescentar um gráfico.

1.2.4 Ainda não terminamos

Na última seção, mostramos como reportar os dados para os modelos mistos. No entanto, fizemos isso apenas porque já sabíamos o que nos aguardava por aqui – e também para

saciar a sua ansiedade, que não aguentava mais nos ver manipulando aqueles dados. Todavia, o fato é que ainda não acabamos. Falta, por fim, fazermos uma discussão sobre a relação entre o nosso modelo e o nosso *design experimental*.

Quando se trata do ajuste de modelos matemáticos, é muito importante pensarmos no nosso experimento. Isso não é uma novidade dos modelos mistos. Quando tratamos da ANOVA também fizemos isso. O *design* determinando a estrutura do nosso modelo.

Para esses experimento em particular, tínhamos os participantes como um fator aninhado (*nested*) na condição experimental; e itens como um fator cruzado (*crossed*) com aquela condição. Assim sendo, quando ajustamos um modelo com interceptos para sujeitos e itens, dissemos para nosso modelo considerar o quanto cada sujeito e cada item variavam, em média, em relação às estimativas de cada grupo, decompondo o fator de erro nos efeitos dos fatores aleatórios.

Contudo, como itens são um fator cruzado, podemos ter um cenário que já discutimos quando tratamos da ANOVA: os itens podem interagir com os níveis das condições, de modo que um item seja mais rápido em uma condição ou mais lento em outra – e vice-versa. A fim de controlar essa variabilidade, devemos, então, acrescentar ao nosso modelo uma inclinação (*slope*) para os itens. O nosso modelo precisará ter, então, a seguinte estrutura de fatores aleatórios (lembre-se da Figura 1.7 e da discussão feita por lá):

1. *para os participantes*: um intercepto, mas não uma inclinação, já que o mesmo participante só via uma condição;

2. *para os itens*: um intercepto e uma inclinação para cada item, já que o mesmo item era visto em ambas as condições.

A Tabela 1.13 mostra esse modelo aplicado aos dados.

Ao aplicarmos esse modelo aos dados que temos, obteremos a seguinte resposta: *"modelo é singular"*. O que isso significa? Para entender mais aprofundadamente isso, você precisará ler Baar et al. (2013) [2]. No entanto, basicamente, isso significa que o seu modelo foi *sobreajustado*, ou seja, que ele é mais complexo do que seus dados podem explicar. Se você reparar bem, agora temos algumas coisas a mais na tabela de efeitos aleatórios: além do intercepto para os itens, temos também uma inclinação. Observe que variância dessa inclinação é muito pequena, muito próxima de zero (ela tem três zeros depois da casa decima – 0.0007539 – enquanto as outras têm no máximo um zero – 0.0424936, 0.0820140 e 0.1325286). É isso que é ser singular, ou seja, ter uma variância que tende a zero. Se isso ocorre, então esse fator não é útil ao modelo, ele não está influenciando em nada e, teoricamente, poderíamos excluí-lo. Além disso, há uma tabela nova ao lado, a tabela de correlação de efeitos aleatórios ("Corr."). Como você pode ver, esse valor é igual a 1, ou seja, 100% da variação explicada pela inclinação já é explicada pela outra covariante.

Não existe consenso, entre os estatísticos, sobre o que fazer nesses casos de singularidade. Vamos seguir aqui a sugestão de Baar et al. (2013): manter a máxima estrutura

Resíduos				
Min	1Q	Mediana	3Q	Max
-2.1894	-0.6736	-0.1192	0.5433	2.2945
Efeitos aleatórios				
Grupos	Nome	Variância	Desvio padrão	Corr.
Participantes	Intercepto	0.0424936	0.20614	
Itens	Intercepto	0.0820140	0.28638	
	Condição Sing	0.0007539	0.02746	1.00
Resíduos		0.1325286	0.36404	
Efeitos fixos				
	Estimativa	Erro padrão	t-valor	
Intercepto	1.00826	0.15909	6.338	
Condição: *sing*	0.30154	0.09922	3.039	

Tabela 1.13 – Modelo linear misto com interceptos para participantes e itens e inclinação para itens em função das condições – estimativas em logaritmo natural de milisegundos com tempo acrescido de 1

possível *antes* de o modelo alcançar ajuste singular. Esse modelo é o que aplicamos antes e já aprendemos a reportar. A única coisa que vamos acrescentar é a informação de que escolhemos aquela estrutura com apenas interceptos, mas não inclinações, porque alcançamos a singularidade. Isso é importante para nosso leitor saber os motivos pelos quais paramos ali.

Ajustou-se aos dados um modelo linear misto, com condição como fator fixo, intercepto para sujeitos e itens como fatores aleatórios – a máxima estrutura antes de o modelo alcançar ajuste singular (Baar et al. (2013) – , de modo que o valor estimado para a condição sing *foi de* 0.30480 *log(seg) com intervalo de confiança (com nível de significância 0.05) de* $0.108, 0.499$*„ indicando que essa condição tende a aumentar o tempo para a primeira fixação em relação à condição* mult*.*

Com isso, podemos nos dar por satisfeitos.

Referências

[1] R Harald Baayen, Douglas J Davidson, and Douglas M Bates. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, 59(4):390–412, 2008.

[2] Dale J Barr, Roger Levy, Christoph Scheepers, and Harry J Tily. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3):255–278, 2013.

[3] Rene Forster, Erica Rodrigues, and Letícia Correa. Quantifier comprehension and the hypothesis of visual saliency. *XX European Conference on eye movements*, 2019.

[4] Jason Osborne. Best practices in data transformation: the overlooked effect of minimal values. *Best Practices in Quantitative Methods*, pages 197–204, 01 2008.

[5] Bodo Winter. *Statistics for linguists: an introduction using R.* Routledge, 2019.

[6] Alain F Zuur, Elena N Ieno, and Chris S Elphick. A protocol for data exploration to avoid common statistical problems. *Methods in ecology and evolution*, 1(1):3–14, 2010.