

Aula 11: Que factores explicam a insatisfação com as condições de trabalho?

Estimação de Modelos de Regressão Logística

Docente: Amílcar Moreira **Data & Hora:** 17/12/2020, 18:00-20:00 **Local:** Edifício F2, Sala 102



Aula 11: Estimação de Modelos de Regressão Logística

E se a nossa variável dependente não é uma variável contínua?

Existe um conjunto de opções ...

Sendo que a escolha do modelo de regressão mais adequado dependerá da natureza da variável dependente e (em alguns casos) da natureza das variáveis dependentes



Aula 11: Estimação de Modelos de Regressão Logística

Variável Dependente	Modelo	Variáveis Independentes
Contínua	Regressão Linear Multipla (OLS)	Todos
Nominal (2 Categorias)	Regressão Logistica Regressão Probit	Todos
Nominal (2+ Categorias)	Regressão Multinomial	Todos
Nominal (2 Categorias)	Regressão Loglinear	Nominais
Ordinal	Ordinal Logit Ordered Probit	Todos
Contagens (contém apenas número inteiros, número finito de valores)	Poisson Regression	Contínuas
Tempo até um evento (Análise de Sobrevivência)	Cox Regression	Todos



Aula 11: Estimação de Modelos de Regressão Logística

• Nesta aula vamos focar-nos num dos modelos de regressão mais comum:

O Modelo de Regressão Logística



Aula 11: Estimação de Modelos de Regressão Logística

Porque é não podemos usar o modelo de regressão linear quando temos uma variável dependente binomial?



Aula 11: Estimação de Modelos de Regressão Logística

Porque é não podemos usar o modelo de regressão linear quando temos uma variável dependente binomial?

 Numa variável dependente contínua os valores distribuem-se à volta de uma linha (a linha de regressão).





Aula 11: Estimação de Modelos de Regressão Logística

Porque é não podemos usar o modelo de regressão linear quando temos uma variável dependente binomial?

- Numa variável dependente contínua os valores distribuem-se à volta de uma linha (a linha de regressão).
- Ora não é isso que acontece com uma variável dependente binária...
- A regressão logística utiliza a curva logística para assim representar a relação entre a variável dependente e as independentes.
- Esta curva assume que a partir de um dado valor da IV, o valor da DV não se altera



Fonte: https://smolski.github.io/livroavancado/reglog.html



Aula 11: Estimação de Modelos de Regressão Logística

Modelo de regressão linear (simples)

$$\gamma = \beta_0 + \beta_1 X + E$$

- γ Variável Dependente
- β_0 Constante (i.e., valor de Y quando X = 0)
- $\beta_1 X \qquad \begin{array}{l} \mbox{Coeficiente Beta da variável X (mede o efeito de uma} \\ \mbox{alteração unitária de X sobre o valor médio da variável Y,} \\ \mbox{quando todas as outras variáveis estão fixas)} \end{array}$
- *E* Erro aleatório ou estocástico (reflete a influência de outros factores no no comportamento da variável Y que não podem ser explicadas linearmente pelo comportamento da variável X)

Modelo de regressão logística

$$ln\left(\frac{p}{1-p}\right) = b_0 + b_1 X_1 + \dots + b_n X_n$$

$ln\left(\frac{p}{1-p}\right)$ Logaritmo natural das chances (odds) de *p* acontecer

- p Probabilidade de um evento acontecer
- 1-p Probabilidade de um evento não acontecer
 - b₀ Constante
- b₁X₁ Coeficiente Beta da variável X (mede o efeito de uma alteração unitária de X sobre o logaritmo das odds da variável dependente)



Aula 11: Estimação de Modelos de Regressão Logística

Implementação do Modelo de Regressão Logística

- 1. Criar a variável dependente
- 2. Inspecionar a variável dependente
- 3. Explorar as relações entre variáveis
- 4. Definir o modelo de regressão logística
- 5. Interpretar os rseultados do modelo de regressão logística
- 6. Validar pressupostos do modelo de regressão logística



Aula 11: Estimação de Modelos de Regressão Logística

- Objectivo:
 - Recodificar a variável que mede a satisfação com as condições de trabalho ('workcond_sat'), numa variável dummy em que 1 identifica pessoas que responderam 'Very satisfied' (1) ou 'Satisfied' (2).



 (\mathbf{C})

Aula 11: Estimação de Modelos de Regressão Logística

Criar a variável dependente

- Selecionar 'Transformar'/'Recodificar A em variáveis diferentes'
- Selecionar a variável 'workcond_sat'... B
 - ... e colocar na caixa da 'variável de entrada'



rquivo	Editar	Visualizar	<u>D</u> ados	Transformar	Analisar	<u>G</u> ráficos	Utilitários	E <u>x</u> tensões	Janela	Ajuda			
		⇒ Щ		<u> 1</u>		μ							
			ta Rec	codificar em vari	áveis diferent	95					~	Visível: 2	25 de 25 vari
	&	departme 🛷 nt			aveis unerent	Variável (de entrada > V	ariável de saí	da: _Va	riával da saí	~eh	🗞 anxiety	🗞 fatigue
1		6	j n	earing kip	-				No	me:	ua	1	
2		5	- 2 bi	ack								1	
3		2	— 👗 m	nuscle_upper		CJ			Ró	itu <u>l</u> o:		1	
4		1	💑 m	nuscle_lower								1	
5		2	hi 🔔 in	eadache viur/						Alton		1	
6		4	— 👗 ai	nxiety						Altera	1	1	
		1	🗕 💑 fa	atigue								1	
0				usent_m								2	
10				orkcond_sat		Valores	antigos e nov					1	
11	_	2	y 1	_wage2								1	
12		- 1	e e	valuation2	-	Se (C	ondição de se	eleção de cas	o opcional)			. 1	
13	_	1			ОК	Colar	Reconfigura	ar Cancel	ar /	Ajuda		1	
14		4	40.5	9		1						1	
15		2	40.0	5 1		1	1	1	1	- 1		1 1	
_	4									_			
/isualiz	ação de	dados Visua	lização de	variável									

11



(F)

Aula 11: Estimação de Modelos de Regressão Logística

- Selecionar 'Transformar'/'Recodificar em variáveis diferentes'
- Selecionar a variável 'workcond_sat'...
 - ... e colocar na caixa da 'variável de entrada'
- Definir o nome da nova variável ('satisfied') e o rótulo da variável ('Satisfied')
- Selecionar o botão 'Alterar'
- Selecionar 'Valores antigo e novo'





Aula 11: Estimação de Modelos de Regressão Logística

- Vamos definir que ao valor '1' e '2' na variável 'workcond_sat'
- Corresponde o valor '1', na nova variável ('satisfied')
- Selecionar 'Incluir'

	Arquivo	<u>E</u> ditar	<u>V</u> isualizar	<u>D</u> ados	<u>T</u> ransformar	<u>A</u> nalisar	<u>G</u> ráficos	<u>U</u> tilitários	E <u>x</u> tensões	Janela	Ajuda			
D			Recodifica	r em variáv	veis diferentes: va	alores antigo	s e novos					×	E do OE I	voriávaja
	1:	455	Valor antigo)				o valor alor: 1					ar	variaveis Vä
_	1						00	misso no sis	tema					
\mathbf{n}	2		Omisso	no sist <u>e</u> m	3		00	o <u>p</u> iar valor(e	s) antigo(s)					
Ŀ	3		Omisso	para o sist	ema ou usuário			Antino	> Maure:					
	4		O Amplitud	e:				Anugo	> 1\0\0.					_
	5													
	7		a <u>t</u> é				Inc	luir	N					
	8	6			(O até o valor:		Alte	erar	y					
	9	9			to ale o valor.		Re <u>m</u>	iover	_					
	10		Intervalo	valor até o	MAIS ALTO:									
	11							🔲 As variáve	is de saída sã	io sequênc	cias de caracteres	argura: 8		
	12		O Todos os	s outros va	lores			Converter	sequências d	le caractere	es numéricas em númer	ros ('5'->5)		
	13		0.7											
	14						Continua <u>C</u> ontinua	Cancelar	Ajuda					
	15	4												



Aula 11: Estimação de Modelos de Regressão Logística

- Vamos definir que ao valor '1' e '2' G na variável 'workcond_sat'
- Corresponde o valor '1', na nova variável ('satisfied')
- Selecionar 'Incluir'
- Vamos definir que 'Todos os outros valores' na variável 'workcond_sat'
- Corresponde o valor '0', na nova variável ('satisfied')
- Selecionar 'Incluir'

	🕼 *AD0	GRH_BD	_Aula11.sav [Con	juntodeDa	idos1] - Editor de	e dados do II	BM SPSS Sta	tistics						_		×
	Arquivo	<u>E</u> dita	r <u>V</u> isualizar	<u>D</u> ados	<u>T</u> ransformar	<u>A</u> nalisar	<u>G</u> ráficos	<u>U</u> tilitários	E <u>x</u> tensões	Janela	Ajuda					
H	1:		Recodifica	r em variáv	veis diferentes: v	alores antigo	s e novos	H 💻						×	5 de 25	variáveis
–	1	4	Valor antigo © Valor:					vo valor /a <u>l</u> or: 0							ar	VE
	2		O Omisso	no sist <u>e</u> ma para o sist	a tema ou usuário)	00	Co <u>p</u> iar valor(e	s) antigo(s)							
\mathbf{O}	4 5 6 7		Amplitud até	e:			Alt	A <u>n</u> tigo Lowe cluir erar	> Novo: st thru 2> 1							
	8 9 10 11		 ○ Intervalo, ○ Intervalo, 	, MAIS BAIX , valor até c	(O até o valor: o MAIS ALTO:		Rer	nover		w			1			
K	12 13 14	O	Todos os	s outros va	lores		(<u>C</u> ontinua	Converter	sequências of Ajuda	ao sequend de caractero	cias de car es nu <u>m</u> ério	acteres :as em nún	Largura: 1eros ('5'->5	5)		
L	15 Visualiz	1 zação d	e dados Visual	ização de v	variável			***								
									O processad	or do IBM S	PSS Statis	tics está p	ronto	Unicode	ON	



Aula 11: Estimação de Modelos de Regressão Logística

- Vamos definir que ao valor '1' e '2' na variável 'workcond_sat'
- Corresponde o valor '1', na nova variável ('satisfied')
- Selecionar 'Incluir'
- Vamos definir que 'Todos os outros valores' na variável 'workcond_sat'
- Corresponde o valor '0', na nova variável ('satisfied')
- Selecionar 'Incluir'
- Selecionar 'Continuar' / 'OK'

Valor antigo Valor antigo Valor: Valor: Omisso no sistema Omisso no sistema Omisso no sistema Omisso para o sistema ou usuário Amplitude: Sector Amplitude: Image: Sector Image: Sector <t< th=""><th>Valor antigo Novo valor Valor: Omisso no sistema Omisso no sistema Omisso no sistema Omisso para o sistema ou usuário Antigo -> Novo: Amplitude: Lowest thru 2-> 1 Incluir Listerar Intervalo, MAIS BAIXO até o valor: Lowest thru 2-> 1 Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres numéricas em números (f5->5)</th><th>Ar</th><th>rquivo <u>I</u></th><th>Editar</th><th>Visualizar</th><th>Dados</th><th>Transform</th><th>nar <u>A</u>nal</th><th>lisar <u>G</u>ra antigos e no</th><th>áficos <u>U</u>ti I ovos</th><th>itários</th><th>E<u>x</u>tensões</th><th>Janela</th><th>Ajuda</th><th></th><th></th><th>×</th><th></th></t<>	Valor antigo Novo valor Valor: Omisso no sistema Omisso no sistema Omisso no sistema Omisso para o sistema ou usuário Antigo -> Novo: Amplitude: Lowest thru 2-> 1 Incluir Listerar Intervalo, MAIS BAIXO até o valor: Lowest thru 2-> 1 Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres numéricas em números (f5->5)	Ar	rquivo <u>I</u>	Editar	Visualizar	Dados	Transform	nar <u>A</u> nal	lisar <u>G</u> ra antigos e no	áficos <u>U</u> ti I ovos	itários	E <u>x</u> tensões	Janela	Ajuda			×	
1 2 0 Omisso no sistema 3 ○ Omisso para o sistema ou usuário 4 ○ Amplitude: 5 6 até 7 8 0 Intervalo, MAIS BAIXO até o valor: 9 10 10 11 12 13 14 0 Omisso no sistema 0 Omisso para o sistema ou usuário Antigo> Novo: Lowest thru 2> 1 Lowest thru 2> 1 ELSE> 0 Atterar Remover Incluir As variáveis de saída são sequências de caracteres Largura: 8 Todos os outros valores Converter sequências de caracteres numéricas em números (5'->5)	 Omisso no sistema Omisso no sistema Omisso para o sistema ou usuário Amplitude: Afigo> Novo: Lowest thru 2> 1 ELSE> 0 Intervalo, MAIS BAIXO até o valor: Intervalo, valor até o MAIS ALTO: Tgdos os outros valores Tgdos os outros valores As variáveis de saída são sequências de caracteres Largura: 8 Ajuda Ajuda 				Valor antigo)				Novo valo Va <u>l</u> or:	r							5 de 25 va
2 Omisso no sistema 3 Omisso para o sistema ou usuário 4 Amplitude: 5 Induir 6 até até Incluir 0 Intervalo, MAIS BAIXO até o valor: 9 Intervalo, valor até o MAIS ALTO: 10 Intervalo, valor até o MAIS ALTO: 11 As variáveis de saída são sequências de caracteres Largura: 8 0 Todos os outros valores 13 Converter sequências de caracteres numéricas em números (5'->5)	 Omisso no sistema Omisso para o sistema ou usuário Amplitude: Amplitude: Lowest thru 2 -> 1 ELSE -> 0 Intervalo, MAIS BAIXO até o valor: Intervalo, valor até o MAIS ALTO: Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Todos os outros valores Converter sequências de caracteres Largura: Continuar Cancelar Ajuda Visualização de variável 		1							Omise	o no siste	ma						38
3 O Omisso para o sistema ou usuário 4 5 6 até 7 até 7 0 Intervalo, MAIS BAIXO até o valor: 9 10 10 11 12 o Intervalo, valor até o MAIS ALTO: 11 12 o Intervalo, valor até o MAIS ALTO: 11 12 o Intervalo, valor até o MAIS ALTO: 11 12 13 14 Antigo -> Novo: Lowest thru 2> 1 Lowest thru 2> 1 LSE -> 0 Incluir Atterar Remover Incluir Atterar Converter sequências de caracteres Largura: 8 Conterter sequências de caracteres numéricas em números ('5->5) 13	Antigo> Novo: Amplitude: até até Incluir Alterar © Intervalo, MAIS BAIXO até o valor: © Intervalo, Valor até o MAIS ALTO: © Intervalo, valor até o MAIS ALTO: © Todos os outros valores Antigo> Novo: Lowest thru 2> 1 ELSE> 0 41.00 41.00 40.90 41.00 40.90 41.00 40.90 41.00 40.90 41.00 40.90 41.00 40.90 40.90 41.00 40.90		2		O Omisso	no sist <u>e</u> m	a			© Co <u>p</u> ia	valor(es)	antigo(s)						39
4 O Amplitude: 5 6 até 7 8 0 Intervalo, MAIS BAIXO até o valor: 9 10 0 Intervalo, valor até o MAIS ALTO: 11 12 13 14 Continuar As variáveis de saída são sequências de caracteres Largura: 8 0 Intervalo, valor até o MAIS ALTO: Continuar Continuar As variáveis de saída são sequências de caracteres Largura: 8 13 14	O Amplitude: 30.0 até Incluir Atterar ELSE -> 0 O Intervalo, MAIS BAIXO até o valor: 41.0 O Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: Ø Todos os outros valores Contentar sequências de caracteres numéricas em números (5->5) O Intervalo, Visualização de variável As variáveis	_	3		O Omisso	para o sis	tema ou usi	uário			Antigo	> Novo:						40
6 até 7 até 8 Incluir 8 Incluir 10 Intervalo, MAIS BAIXO até o valor: 9 Intervalo, valor até o MAIS ALTO: 11 As variáveis de saída são sequências de caracteres Largura: 8 Todos os outros valores 12 Todos os outros valores 13 Converter sequências de caracteres numéricas em números (5'->5) 14 Ajuda	até até Incluir Alterar Remover Incluir Alterar Remover ELSE -> 0 41.0 38.9 41.7 40.9 41.0 40.9 40.6 40.5 40.	_	5		O Amplitud	le:					Lowest	thru 2> 1						38
7 8 9 10 0 Intervalo, MAIS BAIXO até o valor: 0 Intervalo, valor até o MAIS ALTO: 11 12 0 Todos os outros valores 13 14	ade alterar C Intervalo, MAIS BAIXO até o valor: C Intervalo, valor até o MAIS ALTO: Todos os outros valores As variáveis de saída são sequências de caracteres Largura: C Converter sequências de caracteres numéricas em números (5:->5) 39.9 Continuar Cancelar Ajuda 40.6 41.0 40.6 41.0 40.6 41.0 40.6 41.0 40.6 41.0 40.6 41.0 40.6 41.0 40.6 40		6		otó					Incluir	ELSE	> <mark>0</mark>						41
8 Intervalo, MAIS BAIXO até o valor: 9 Intervalo, valor até o MAIS ALTO: 11 As variáveis de saída são sequências de caracteres Largura: 8 Converter sequências de caracteres numéricas em números ('5'->5) 13 Continuar Cancelar Ajuda	Atterar 41.7 O Intervalo, MAIS BAIXO até o valor: 40.9 O Intervalo, valor até o MAIS ALTO: 41.0 Image: Converter sequências de caracteres largura: 8 Image: Converter sequências de caracteres numéricas em números ('5'->5) 39.9 Image: Continuar Cancelar Ajuda 40.6		7		ale					inciuii								38
9 Image: Second and or value. 10 Intervalo, valor até o MAIS ALTO: 11 As variáveis de saída são sequências de caracteres Largura: 12 Todos os outros valores 13 Converter sequências de caracteres numéricas em números ('5'->5) 14 Continuar	Intervalo, mino Druko ate o valor. 40.9 Intervalo, mino Druko ate o valor. 40.9 Intervalo, valor até o MAIS ALTO: 41.0 Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: 40.9 Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: 40.9 Intervalo, valor até o MAIS ALTO: Converter sequências de caracteres numéricas em números (15->5) 39.9 Intervalo, minor Cancelar Ajuda 40.6 Intervalor Intervalor 40.6 Intervalor Intervalor 40.6 Intervalor Intervalor 40.6		8				(O até o valu			Alterar								41
10 11 12 13 14 10 10 11 12 13 14 10 11 13 14 14 10 13 14 14 14 14 14 10 13 14 14 13 14 14	Intervalo, valor até o MAIS ALTO: As variáveis de saída são sequências de caracteres Largura: As variáveis de saída são sequências de caracteres Largura: As variáveis de saída são sequências de caracteres Largura: Todos os outros valores Converter sequências de caracteres numéricas em números ('5'->5) 39.9 Continuar Cancelar Ajuda 40.6 41.1 40		9			, MAIO DAI				Re <u>m</u> over								40
11 12 13 14 14 14 As variáveis de saída são sequências de caracteres Largura: 8 Converter sequências de caracteres numéricas em números ('5'->5) Continuar Cancelar Ajuda	As variáveis de saída são sequências de caracteres Largura: As variáveis de saída são sequências de caracteres Largura: As variáveis de saída são sequências de caracteres Largura: As variáveis de saída são sequências de caracteres Largura: As variáveis de saída são sequências de caracteres Largura: Attracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas em números ('5'->5) As variáveis de saída são sequências de caracteres numéricas		10		O Intervalo	valor até).										41
12 13 14 0 12 13 14 0 14 0 12 Converter sequências de caracteres numéricas em números ('5'->5) 14 Image: Continuar Cancelar Ajuda	Todos os outros valores Converter sequências de caracteres numéricas em números ('5'->5) 41.1 39.9 40.5 40.6		11			, 10.01 0.0				A I	variáveis	de saída sâ	io sequênc	cias de car	acteres	Largura:	8	40
13 14 <u>Continuar</u> Cancelar Ajuda	39.9 40.5 40.0 acão de dados Visualização de variável		12		Todos os	s outros va	lores				onverter s	equências d	e caractere	es numério	as em núi	- meros ('5'->	5)	41
14 Continuar Cancelar Ajuda	Continuar Cancelar Ajuda 40.5 40.6		13		0.1									<u></u>				39
	acão de dados Visualização de variável		14						<u></u>	ontinuar	ancelar	Ajuda						40
15	ação de dados Visualização de variável		15	4														40



Aula 11: Estimação de Modelos de Regressão Logística

Criar a variável dependente

 Vamos verificar se a recodificação foi bem sucedida

ta *AD	GRH_B	D_Aula11.s	av [Cor	njuntodeDao	dos1] - Editor de	e dados do IBM	SPSS Statistics					—		×
Arquivo	Edita	ar <u>V</u> isua	əlizar	Dados	Transformar	<u>A</u> nalisar (gráficos <u>U</u> til	itários E <u>x</u> tensões	Janela	Ajuda				
				5	🔁 🔛		P 44			•				
1 : satisf	ied		.0									Visível:	26 de 26 v	ariáveis
		🥙 absent	_nr 🖬	workcond _sat	I evaluation	✓ y_wage2	evaluation 2	\lambda 윩 satisfied	var	var	var	var	var	
1			9	3	49.08	17529.60	55.17	.00						
2			10	2	52.58	24843.10	51.75	1.00						
3			15	2	53.80	29874.72	54.79	1.00						
4			10	2	50.42	20994.46	58.30	1.00						
5			11	4	48.43	10044.01	49.89	.00						
6			10	2	54.34	32321.53	52.74	1.00						
7			6	3	49.65	18757.44	47.60	.00						
8			10	4	46.84	9106.36	52.65	.00						
9			8	2	50.76	20432.69	48.48	1.00						
10			17	3	53.33	29327.90	52.40	.00						
11			14	2	50.59	18223.72	47.21	1.00						
12			7	2	54.64	29668.47	48.39	1.00						
13			13	3	56.64	35658.63	50.19	.00						
14			7	3	55.36	35733.01	40.86	.00						
15		0	11	3	56.09	36160 52	45 36	nn						-
Visuali	zação d	le dados	Visual	lização de v	ariável			O processad	lor do IBM SF	2SS Statistic	s está pronto	Unicode	ON	



Aula 11: Estimação de Modelos de Regressão Logística

Implementação do Modelo de Regressão Logística

2. Inspecionar a variável dependente



Aula 11: Estimação de Modelos de Regressão Logística

- Objectivo:
 - Perceber como se distribuem as observações na variável dependente ('satisfied').

• Neste caso, através de um Gráfico de Barras...



Aula 11: Estimação de Modelos de Regressão Logística

Gráfico de Barras

 Selecionar 'Gráficos' / 'Construtor de Gráfico'



ta Adgrh_i	3D_Aula5.sav [Co	onjuntodeDados	s1] - Editor de d	ados do IBM S		.5			
Arquivo <u>E</u> e	ditar <u>V</u> isualiz	ar <u>D</u> ados	<u>T</u> ransformar	<u>A</u> nzusar (<u>G</u> ráficos	<u>U</u> tilitários	E <u>x</u> tensões	Janela	Ajuda
				┹=	🎁 <u>C</u> onstru	itor de Gráfic	:0		
		• •	-		Seletor	de modelo d	le tabelas de <u>o</u>	gráficos,	
					Caixas	ue unajogo i	egadas		
	¢¢ id	🛷 year	🗞 sex	🛷 age 🖢	🍑 a ec	lucation	✓ y_w	age	art_yr
1	1	2018	1	34	ISCED1		1588	83.75289	2013
2	2	2018	2	35	ISCED1		2108	82.22292	2008
3	3	2018	2	42	ISCED4				2005



Aula 11: Estimação de Modelos de Regressão Logística

- Selecione 'Barras'
- Selecionar (com duplo-clique) C o Gráfico de Barras (simples)

Variáveis:	A visualização do gráfico usa dados de exemplo			
		L¢1	Propriedades do elemento Aparência do Gráfico Opções	
Id Image: Sex <	Arraste um gráfico da Galeria aqui para usá-lo como seu ponto de início OU Clique na guia Elementos Básicos para construir um elemento de gráfico por elemento		Editar propriedades de:	
Área Setor/Polar Dispersão/Ponto Histograma Alto-Baixo: Boxplot Eixos duplos				
	 ✓ id ✓ year ✓ sex age e ducation ✓ y_wage ✓ start_yr The department ← hrs_week ← hearing ✓ Nenhuma categoria (variável de escala) Galeria Elementos básicos II Escolha entre: Favoritos Barra Linha Área Setor/Polar Dispersão/Ponto Histograma Alto-Baixo: Boxplot Eixos duplos 	✓ Id ✓ year Sex age education ✓ yage ✓ start_yr department hrs_week hearing ✓ Venthuma categoria	✓ vid ✓ year Sex age deducation ✓ y.wage Start_vr department hrs_week hearing ✓ Vertice Clique na guia Elementos Básicos para construir um elemento de gráfico por elemento Galeria Elementos básicos D de Conto Titulos/Notas de rodapé Esgolha entre: Favoritos Barra Linha Área Setor/Polar Dispersão/Ponto Hosgrama Ato-Baixa Borgola Exortos Borgola OK Colar Reconfigurar Cancelar Araste um gráfico da Galeria aqui para usá-lo como seu ponto de início OU Clique na guia Elementos Básicos para construir um elemento de gráfico por elemento Setor/Polar Dispersão/Ponto Histograma Ato-Baixa Exos duplos OK Colar Reconfigurar Cancelar Ajuda	Image: Sex (a) Arraste um gráfico da Galeria anul para usá-lo como seu ponto de início Sex (a) OU Y (wage) Cilque na guia Elementos Básicos para construir um elemento de gráfico por elemento Menturae categorias (variável de escuela) OU Caterá Elementos básicos D de conto Thulos/Notas de rodapé Escolha entre: Parar atraita I.Linha Arraste Arraste Dispersão/Ponto Histograma Alto-Bário: Borplot Exos duplos OK Cotar



Aula 11: Estimação de Modelos de Regressão Logística

- Selecione 'Barras'
- Selecionar (com duplo-clique) o Gráfico de Barras (simples)
- Selecionar a variável 'satisfied'
 E
- Colocar a variável 'satisfied' no F
 'Eixo X'





Aula 11: Estimação de Modelos de Regressão Logística

- Selecione 'Barras'
- Selecionar (com duplo-clique) o Gráfico de Barras (simples)
- Selecionar a variável 'satisfied'
 E
- Colocar a variável 'satisfied' no F
 'Eixo X'
- Selecionar 'Porcentagem'
- Selecionar 'OK'





Aula 11: Estimação de Modelos de Regressão Logística

- O gráfico é publicado no 'Visualizador de Resultados'
- Menos de metade dos trabalhadores está satisfeito com as condições de trabalho
- Gráfico dever ser complementado com tabela de frequências (ver Aula 5)





Aula 11: Estimação de Modelos de Regressão Logística

Implementação do Modelo de Regressão Logística

3. Explorar as relações entre variáveis



Aula 11: Estimação de Modelos de Regressão Logística

- Objectivo:
 - Perceber como se relaciona a variável dependente ('satisfied') com a variável independente 'sex' que queremos incluir no nosso modelo de regressão

... neste caso, através de um Gráfico de Barras Sobrepostas



(A)

Aula 11: Estimação de Modelos de Regressão Logística

Gráfico de Barras Sobrepostas

 Selecionar 'Gráficos' / 'Construtor de Gráfico'





Aula 11: Estimação de Modelos de Regressão Logística

- Selecionar 'Gráficos' / 'Construtor de Gráfico'
- Selecione 'Barra'
- Duplo-Clique na opção 'Gráfico de Barras Sobrepostas'





Aula 11: Estimação de Modelos de Regressão Logística

Gráfico de Barras Sobrepostas

- Selecionar 'Gráficos' / 'Construtor de Gráfico'
- Selecione 'Barra'
- Duplo-Clique na opção 'Gráfico de Barras Sobrepostas'
- Selecionar a variável 'satisfied'...

... e colocar na caixa do 'Cluster em X'

Exercício: Colocar a variável 'sex' no 'Eixo X'





Aula 11: Estimação de Modelos de Regressão Logística

- Seleccionar 'Barra 1'
- Definir 'Percentagens' como parâmetro de comparação





Aula 11: Estimação de Modelos de Regressão Logística

- Seleccionar 'Barra 1'
- Definir 'Porcentagem' como parâmetro de comparação
- Selecionar 'Configurar Parâmetros'





Aula 11: Estimação de Modelos de Regressão Logística

- Seleccionar 'Barra 1'
- Definir 'Porcentagem' como parâmetro de comparação
- Selecionar 'Configurar Parâmetros'
- Selecionar o denominador 'Total para cada categoria do eixo X'
- Selecionar 'Continuar' / 'OK'





Aula 11: Estimação de Modelos de Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Não se notam diferenças significativas na satisfação com as condições de trabalho em função do sexo...
- Gráfico deve ser complementado com tabela de dupla entrada para estas variáveis (ver Aula 6)





Aula 11: Estimação de Modelos de Regressão Logística

- Objectivo:
 - Comparar a idade média dos trabalhadores que estão satisfeitos com as condições de trabalho com a idade média dos trabalhadores que não estão satisfeitos.

... neste caso, através de um Gráfico de 'Barra de Erro Simples'



(A)

Aula 11: Estimação de Modelos de Regressão Logística

Gráfico de Barra de Erro Simples

 Selecionar 'Gráficos' / 'Construtor de Gráfico'





Aula 11: Estimação de Modelos de Regressão Logística

Gráfico de Barra de Erro Simples

- Selecionar 'Gráficos' / 'Construtor de Gráfico'
- Selecione 'Barra'
- Duplo-Clique na opção 'Gráfico Barra de Erro Simples'





Aula 11: Estimação de Modelos de Regressão Logística

Gráfico de Barra de Erro Simples

- Selecionar 'Gráficos' / 'Construtor de Gráfico'
- Selecione 'Barra'
- Duplo-Clique na opção 'Gráfico Barra de Erro Simples'
- Selecionar a variável 'satisfied'...

... e colocar na caixa do 'Eixo X'

Exercício: Colocar a variável 'age' no 'Eixo Y'





Construtor de gráfico

Aula 11: Estimação de Modelos de Regressão Logística

Gráfico de Barra de Erro Simples

 Escolher a 'Média' como estatística a ser representada no gráfico

F	<u>V</u> ariáveis:	A visualização do gráfico usa dados de exemplo	Propriedades do elemento Aparência do Gráfico Opções
F		Barra de Erro Simples Média de age por Satisfied	Editar propriedades do etementa do chained opções Editar propriedades de: Ponto1 X-Eixo1 (Ponto1) Y-Eixo1 (Ponto1) Y-Eixo1 (Ponto1) Yana 1 Estatísticas Variável: Ø age Estatística: Média Valor Média Mediana Mediana Mediana Máximo N válido Matino Multiplicador: 2
	Linna Área Setor/Polar Dispersão/Ponto Histograma Alto-Baixo: Boxplot Eixos duplos	Image: Colar Reconfigurar Cancelar Ajuda	Empilhar valores idênticos Exibir linhas de projeção verticais entre pontos Linhas de Ajuste Lineares Total Subgrupos

X



Aula 11: Estimação de Modelos de Regressão Logística

Gráfico de Barra de Erro Simples

- Escolher a 'Média' como estatística a ser representada no gráfico
- Desligar a opção 'Exibir Barra de Erros'

• Clicar 'OK'





Aula 11: Estimação de Modelos de Regressão Logística

Gráfico de Barra de Erro Simples

• O resultado é publicado no 'Visualizador de Resultados'

- Não se notam diferenças significativas em termos da idade média relativamente à sua satisfação com as condições de trabalho ...
- Gráfico deve ser complementado com tabela de comparação de médias (ver Aula 6)





Aula 11: Estimação de Modelos de Regressão Logística

- Objectivo:
 - Perceber como se relaciona a variável dependente ('satisfied') com a variável independente 'education' que queremos incluir no nosso modelo de regressão

... neste caso, através de um Gráfico de Barras Agrupadas



Aula 11: Estimação de Modelos de Regressão Logística

Gráfico de Barras Agrupadas

 Selecionar 'Gráficos' / 'Construtor de Gráfico'







Aula 11: Estimação de Modelos de Regressão Logística

- Selecionar 'Gráficos' / 'Construtor de Gráfico'
- Selecione 'Barra'
- Duplo-Clique na opção 'Gráfico de Barras Agrupadas'





Aula 11: Estimação de Modelos de Regressão Logística

Gráfico de Barras Agrupadas

- Selecionar 'Gráficos' / 'Construtor de Gráfico'
- Selecione 'Barra'
- Duplo-Clique na opção 'Gráfico de Barras Agrupadas'
- Selecionar a variável 'satisfied'...
 - ... e colocar na caixa do 'Cluster em X'

Exercício: Colocar a variável 'education' no 'Eixo X'





Aula 11: Estimação de Modelos de Regressão Logística

- Seleccionar 'Barra 1'
- Definir 'Porcentagem' como parâmetro de comparação





Aula 11: Estimação de Modelos de Regressão Logística

- Seleccionar 'Barra 1'
- Definir 'Porcentagem' como parâmetro de comparação
- Selecionar 'Configurar parâmetros'





Aula 11: Estimação de Modelos de Regressão Logística

- Seleccionar 'Barra 1'
- Definir 'Porcentagem' como parâmetro de comparação
- Selecionar 'Configurar parâmetros'
- Selecionar o denominador 'Total para cada categoria do eixo X'
- Selecionar 'Continuar' / 'OK'





Aula 11: Estimação de Modelos de Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Podemos perceber que a percentagem de trabalhadores satisfeitos com as condições do trabalho aumenta em função da educação
- Deve ser complementado com uma tabela de dupla entrada (ver Aula 6)





Aula 11: Estimação de Modelos de Regressão Logística

Implementação do Modelo de Regressão Logística

4. Definir o modelo de regressão logística



Aula 11: Estimação de Modelos de Regressão Logística

- Objectivo:
 - Estimar um modelo de regressão logística que explique em que medida a probabilidade de se sentir satisfeito com as condições de trabalho varia em função do sexo, idade e educação.



(A**)**

Aula 11: Estimação de Modelos de Regressão Logística

Regressão Logística

 Selecionar 'Analisar'/ 'Regressão' / 'Logística binária'

	1				2 24-5						12		
Arquivo	Editar	Visualiza	ir <u>D</u> ados <u>I</u>	ansformar	Analisar	Graticos	Utilitarios	Extensoe	s Janel	i Aju	ida		
			5 2		Relati <u>E</u> stati	órios stica Descri	tiva	•	A 14				
12 : y_wa	ge2	296	68.4665929268	8	Estatí	sticas <u>B</u> aye	siana	•					
	and a	absent_nr	workcond _sat	evaluation	Ta <u>b</u> el Comr	as barar Médias		*	var		var	var	
1		9	3	49.08	Model	lo Linear Ge	ral						
2		10	2	52.58	Model	los lineares	generalizados						
3		15	2	53.80	Model		generalizados						
4		10	2	50.42	Woder	103 11113105							
5		11	4	48.43	Deere	acional						1200	
6		10	2	54.34	Regre	essao			Model <u>a</u>	agem Li	near Auto	mática	
7		6	3	49.65	L <u>og</u> III	near			Linear				
8		10	4	46.84	Rede	s neurais		1	🗾 <u>C</u> urva	de estin	nação		
9		8	2	50.76	Class	lificar		,	🔣 Quadr	ado <u>s</u> Mí	nimos Pa	rciais	
10		17	3	53.33	Re <u>d</u> u	ção de dime	nsão	•	L ogísti	ica hinái	ria		
11		14	2	50,59	Esc <u>a</u> l	а		•	Logion	ou oniu			



Aula 11: Estimação de Modelos de Regressão Logística

- Selecionar 'Analisar'/ 'Regressão' / 'Logística binária'
- Selecionar a variável 'satisfied'
- Colocar na caixa 'Dependente'
- **Exercício**: Colocar as variáveis 'sex', 'age' e 'education' na caixa 'Covariáveis'





Aula 11: Estimação de Modelos de Regressão Logística

- Selecionar 'Analisar'/ 'Regressão' / 'Logística binária'
- Selecionar a variável 'satisfied'
- Colocar na caixa 'Dependente'
- **Exercício**: Colocar as variáveis 'sex', 'age' e 'education' na caixa 'Covariáveis'
- Clicar no botão 'Categórico'

đ	a *ADG	RH_BD_	Aula11.sa	iv [Con ilizar	juntodeDao Dados	dos1] - Editor de	e dados do I Analisar	BM SPSS Stat	tistics Utilitários	Extensões	Janela	Ajuda		_		×
(10,		-	• =	ч	#1 💻	1			_			
12	2 : y_wa	ge2		29668	46659292	ta Regressão	logística					×	<	Visível:	26 de 26 variá	vei
			🛷 id		🜮 year	id.			Dependente:	- 1- 5 - 13		Cate <u>q</u> órico	y_wage	departme nt		2
	1			1	2018.00	sear		Elloco 1	atistied [1 de 1 -	satistiedj		Salvar	22153.05	6	38.85	;
	2			2	2018.00	💑 sex		Antori		G	Trávima	Opções	27809.35	5	39.18	;
	3			3	2018.00	age		Anten		Ľ	TOXITTO	Estilo	29509.03	2	40.24	ł
	4			4	2018.00	start vr	'		Blo <u>c</u> o 1 de	1		Bootstrap	21500.50	1	38.62	2
	5			5	2018.00	🛷 start_age	•		sex				16786.81	2	38.86	j -
	6			6	2018.00	experient	ce		education				31153.80	4	41.00	1
	7			7	2018.00	y_wage	ont	>a*b>					20867.09	1	38.90	4
	8			8	2018.00	hrs_wee	k						4899.42	3	41.70	4
	9			9	2018.00	💑 hearing		Método): Inserir		Ŧ		20882.25	1	40.95	4
	10	_		10	2018.00	and skin			Variável de sel	ecão:			29631.50	5	41.07	4
	11			11	2010.00	muscle i	upper		Vanaver de Ser	cyao.	Regra		2090.40		40.01	
	12	_		12	2018.00	a muscle	ower				rtogra		1361.65	1	39.90	-
	14	_		14	2018.00		ж	Colar	Reconfigura	Cancel	ar 🛛	Ajuda	30477 37	4	40.59	
	15			15	2018.00	1		48	3 20	03 00	33.00	15.00	32136.62	2	40.05	
•	/isualiza	ação de	e dados	Visuali	zação de v	ariável										
										O processad	or do IBM S	PSS Statistics	está pronto	Unicode	:ON	ĺ



E

(F)

Aula 11: Estimação de Modelos de Regressão Logística

- Selecionar a variável 'sex'
- Colocar na caixa 'Covariáveis categóricas'

vrquivo į	Editar ⊻isu	alizar <u>D</u> ado	s <u>T</u> ransformar	<u>A</u> nalisar	Gráficos	Utilitários	E <u>x</u> tensões	Janela	Ajuda			
2					μ		1	10		_		
12 : y_wage	2	29668.46659	292 🤷 Regressão	o logística						×	Visível:	26 de 26 varia
	🧬 id		Regressão	o logística: defir	nir variávei	s de categoria				× _wage	e 💰 departme nt	✓ hrs_wee
1						Covariáveis ca	tegóricas:			2153.0)5 6	38.8
2		2 201	B.O Sex							7809.3	35 5	39.1
3		3 201	B.0 educatio	on						9509.0	03 2	40.2
4		4 201	B.O							1500.	50 1	38.6
5		5 201	B.0							6786.8	31 2	38.8
6		6 201	B. O		$\mathbf{\mathbf{x}}$					1153.8	30 4	41.0
7		7 201	B.0		(F)					0867.)9 1	38.9
8		8 201	B.0		U	- Alterar Contra	iete			4899.4	12 3	41.7
9		9 201	B.0			Contractor	Inte	dioodor	Altoror	0882.2	25 1	40.9
10		10 201	B. O			Co <u>n</u> traste.		luicador	Alteral	9631.	50 5	41.0
11		11 201	B.0			Categoria de	referencia: @	U <u>i</u> timo	Primeir	2096.4	46 2	40.6
12		12 201	B.0		Continu	arCancelar	Aiuda			8315.	75 1	41.1
13		13 201	B.0		(1361.0	55 1	39.9
14		14 201	B.OC	ок	Colar	<u>R</u> econfigura	r Cancela	r A	juda	30477.3	37 4	40.5
15		15 201	R 00	1 4	18	3 20	00 200	33.00	15.00	32136	² 2	40.0



(E)

(F)

G

(H)

Aula 11: Estimação de Modelos de Regressão Logística

Regressão Logística

- Selecionar a variável 'sex'
- Colocar na caixa 'Covariáveis categóricas'
- Definir a categoria 'male' (valor 1 *) como categoria de referência na variável 'sex'

* neste caso, escolhemos o valor 1 que é o mais baixo (i.e., o primeiro)

• Carregar no botão 'Alterar'

						<u>Grancos</u>		Ežiensues		Ajuda			
12 : y_wa	ge2	29	668.46659292	Regressão	logística					×		Visível:	26 de 26 var
		🛷 id	🛷 year	ta Regressão I	ogística: defi	nir variávei	s de categoria			×	_wage	departme nt	Ins_wee
1		1	2018.0	<u>C</u> ovariáveis:		_	Covariáveis ca	tegóricas:			2153.05	6	38.8
2		2	2018.0	🛷 age			sex(Indicador)			7809.35	5	39.1
3		3	2018.0	education	1						9509.03	2	40.2
4		4	2018.0								1500.50	1	38.6
5		5	2018.0								6786.81	2	38.
6		6	2018.0								1153.80	4	41.(
7		7	2018.0								867.09	1	38.9
8		8	2018.0				- Alterar Contr	acto			9.42	3	41.
9		9	2018.0				Alteral Contr	aste	In dian day of		0882.25	1	40.9
10		10	2018.0				Contraste:		indicador *	Alterar	9631.50	5	41.0
11		11	2018.0				Categoria de	referência:	© U <u>l</u> timo	Primeiro	2096.46	2	40.0
12		12	2018.0			Continua	ar Cancelar	Aiuda		(G)	8315.75	1	41.1
13		13	2018.0			(1361.65	1	39.
14		14	2018.0		Ж	Colar	Reconfigura	Cancel	lar Aj	uda	30477.37	4	40.
15	4	15	2018.0	h 1		48	3 2	003.00	33.00	15.00	32136 62	2	40
			line a Wender	un ri é un l			***						



Aula 11: Estimação de Modelos de Regressão Logística

Regressão Logística

- Selecionar a variável 'sex'
- Colocar na caixa 'Covariáveis categóricas'
- Definir a categoria 'male' (valor 1 *) como categoria de referência na variável 'sex'

* neste caso, escolhemos o valor 1 que é o mais baixo (i.e., o primeiro)

Carregar no botão 'Alterar'

Exercício: Repetir esta operação para a variável 'education'

Selecionar 'Continuar'





 \bigcirc

Aula 11: Estimação de Modelos de Regressão Logística

Regressão Logística

• Clicar no botão 'Opções'

1 9 2 10 3 15 4 10 5 11 6 10 7 6 10 17 9 8 10 17 11 14 12 7	Arquivo E	ditar <u>V</u> is	sualizar	Dados	<u>Transformar</u>	Analisar	<u>Gráficos</u>	<u>U</u> tilitários	E <u>x</u> tensões	Janela	Ajuda				^
12 : y_wage2 29668.46659292 Regressão logística X Visível: 26 de 26 variáv 1 9	a 6				<u>ש</u>		ч	#	1						
Image: State of the state	12 : y_wage2		29668.4	6659292	Regressão I	ogística					×		Visível	26 de 26	variáve
13 13 13 14 7 3 OK Colar Reconfigurar Cancelar Ajuda 15 11 3 56.09 36160.52 45.36 00	1 2 3 4 5 6 7 8 9 10 11 11 12 13 14 15	abse abse	nt_nr 9 9 10 15 10 11 10 6 10 6 10 6 10 8 17 14 7 13 7 11	vorkconc _sat	departme hrs_week hearing skin back muscle_u headache injury anxiety fatigue workcond veuluation y_wage2 veuluation	nt ippper ower sat 12 K	Bloco Anteri >a"b> Método Wétodo	Dependente: Satisfied 1 de 1 Bloco 1 de sex(Cat) age departmen variável de se Reconfigura 45.36	[satisfied] P 1 nt(Cat) leção: ar Cancela	róximo	Categórico Salvar Opções Esuio Booistrap	var	var	var	



Aula 11: Estimação de Modelos de Regressão Logística

- Clicar no botão 'Opções'
- Selecionar a opção 'Qualidade do ajuste de Hosmer-Lemeshow'
- Selecionar a opç 'satisfied'
- Selecionar 'Continuar' / 'OK'

Arquivo	<u>E</u> ditar	Visualizar	<u>D</u> ados	Transformar	<u>A</u> nalisar	<u>G</u> ráficos	<u>U</u> tilitários	E <u>x</u> tensões	Janela	Ajuda			
		🗎 🛄			<u>* 3</u>	μ	AA 💻					-	
12 : y_w	ige2	2966	Regres	são logística: op	ções						×	Visível: 26	de 26 variá
		🛷 id	_ Estatís	ticas e gráficos								departme 🛷	hrs week
			Gra	áficos de classif	icação			Correla	ções de e	stimativas		nt	-
1				euic ob ebchile	a de Hosmi	ar-Lemeshr	114/	Histório	co de iterad	cão		6	38.85
2						er-Lennesine			evn(B)	05 %	`	5	39.18
3		3		tagem por caso	de residuos	5			evb(p);	55 % L	/	2	40.24
4		4	Val	ores discrepant	es no lado d	le fora 2	std. dev					1	38.62
5		5	© To	dos os c <u>a</u> sos								2	38.86
6		6	- Exibir									4	41.00
7		7	. <u>⊚</u> <u>E</u> m	cada passo 🔘	Na ú <u>l</u> tima p	asso						1	38.90
8		8	r Probat	oilidade para Ste	pwise							3	41.70
9		9	Entrad	a: 0 05 Remo	cão: 0 10	1			C	orte de classificação:	0.5	1	40.95
10		10	-	0.00]			Ma	áximo de iterações:	20	5	41.07
11		11	Pres	ervar memória	nara análice	s complexa	e ou grandes	conjuntos de	anten			2	40.61
12		12			para ananoe	is complexe	io ou gran <u>a</u> co	conjuntos de	00000			1	41.16
13		13	V Inclu	iir con <u>s</u> tante no	modelo	\frown						1	39.90
		14				Continua	Cancelar	Ajuda				4	40.59
14							_					0	10.01



Aula 11: Estimação de Modelos de Regressão Logística

Implementação do Modelo de Regressão Logística

5. Interpretar os resultados do modelo de regressão logística



Aula 11: Estimação de Modelos de Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Temos três grupos de resultados
 - Tabelas com informações sobre as variáveis incluídas no modelo
 - Tabelas com os resultados de um modelo de regressão sem variáveis independentes (Bloco 0)
 - Tabelas com os resultados do modelo de regressão que estimamos (<u>Bloco 1</u>)

		Jalizador d	io IBM SPSS Sta	tistics								-	
Arguivo	Editar Visualizar	Dados	Transformar	Inserir F	ormatar	Analisar Gráf	cos Utilitá	rios E	Extensões	Janela	Aiuda		
	-			∼⊒ 🦉			2	1					
	ida Log Regressão logístic Carlo de pr Codificação de pr Codificação de El Bloco 0: Bloco Cim Titulo Cim Tabela de Cim Tabela de Cim Carlo de pr Cim Codificação so Cim Titulo	a ocessame variável d le variáveis Inicial Classifica	ento do caso lependente s categóricas ação	LOGISTIC /METHO /CONTR /CONTR /CRITEF /CRITEF	C REGRE D=ENTE AST (se) AST (edi GOODFI RIA=PIN(0	SSION VARIAE R sex age educ ()=Indicator(1) ucation)=Indica T Cl(95) 0.05) POUT(0.1 gística	LES satisfie ation or(1))) ITERATE	d (20) CL	JT(0.5).				
	🔚 Variáveis i	na equaça não presei	io ntes na equa		Resu	mo de proces	samento d	o caso					
Ē	Bloco 1: Métod	o = Enter		Casos nã	o pondera	dos ^a		V	Porcentager	m			
	Título	Omnibus	do Modelo de	Casos se	lecionado	s Incluído na a	nálise	5000	100	.0			
	Resumo d	do modelo	do modelo di			Casos omis	sos	0		.0			
	Teste de l	Hosmer e l	Lemeshow			Total		5000	100.	.0			
	abela de	Clossifier	icia para test	Casaa nã	io selecior			0		0			
		CIASSIIICA	acão	Casus na	0 00100101	lados		U		.0			
	Variáveis i	na equaçã	ação io	Total	nondaraa	ão estivor em via	ur voja a teho	5000	100. ssificação	.0			
	Uariáveis i	na equaçã	ação io	Total a. Se a para Cod variáv Valor orig .00 1.00	ponderaç o número lificação el deper	ão estiver em vig total de casos. de ndente or interno 0 1	ır, veja a tabe	0 5000 a de cla	100. ssificação	.0			
	└─ û Variáveis i	na equaçã	ação io	Total a. Se a para Cod variáv Valor orig .00 1.00	i ponderação o número lificação el deper inal Val	aaos ão estiver em vig total de casos. de ndente or interno 0 1 3 codificações d	ar, veja a tabe e variávei:	o de cla a de cla s categ	100. ssificação jóricas ão de parân	.0			
	└ () Variáveis i	na equaçã	ação io	Total a. Se a para Cod variáv Valor orig .00 1.00	lificação el deper inal Val	ados ao estiver em vig total de casos. de ndente or interno 0 1 codificações d Frequência	e variávei:	s cates Codificaç (2)	100. ssificação jóricas ão de parân (3)	netro	(4)		
	└ () Variáveis i	na equaçã	ação io	Total a. Se a para Cod variáv Valor orig .00 1.00	ponderação o número el depen inal Val	aados ão estiver em vig total de casos. de ndente or interno 0 1 Codificações d Frequência 1 979	e variávei: (1) (1) (200	s categ codificaç (2)	100. ssificação óricas ão de parân (3) 00 .0	netro 0	(4) .000		
	└ (Î) Variáveis i	na equaçã	ação io	Total a. Se a para Cod variáv Valor orig .00 1.00	ponderação o númerc el deper inal Val	aados ão estiver em vig total de casos. de ndente or interno 0 1 Codificações d Frequência 1 979 1 2 1023	e variávei: (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)	s cates codificaç (2)	100. ssificação fóricas ão de parân (3) 00 .00	netro (000	(4) .000 .000		



(A)

(в)

Aula 11: Estimação de Modelos de Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Estas tabelas dizem-nos:
 - O número de observações que foi integrado no modelo
 - Como está operacionalizada a variável dependente





Aula 11: Estimação de Modelos de Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Estas tabelas dizem-nos:
 - O número de observações que foi integrado no modelo
 - Como está codificada a variável dependente
 - Como estão codificadas a variáveis independentes *
 - * Permíte-nos confirmar se as categorias de referência estão bem definidas





Aula 11: Estimação de Modelos de Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Esta tabela apresenta o resultado de um teste de Chi-Quadrado que determina se, comparado com um modelo sem variáveis independentes, este modelo é estatísticamente significativo.
- Interpretação:
 - 'Sig.' < 0.01 / 0.05 / 0.1 o modelo é estatísticamente significativo
- Neste caso, o nosso modelo é estatísticamente significativo





Aula 11: Estimação de Modelos de Regressão Logística

Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Esta tabela apresenta os resultados de um conjunto de testes, nomedamente o teste R² Nagelkerke, que é um pseudo-R².
- Não pode ser interpretado de forma literal como o R²
- Interpretação:

O model explica aproximadamente 40% da variação na variável dependente





Aula 11: Estimação de Modelos de Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Esta tabela dois indicadores da capacidade preditiva do modelo.
 - <u>Especificidade</u>: A proporção de casos negativos correctamente identificados (verdadeiros negativos)
 - <u>Sensibilidade</u>: A proporção de casos positivos correctamente identificados (verdadeiros positivos)
 - <u>Precisão</u>: A proporção de casos correctamente identificados

n *Saída1 [Documento1] - Visualizador do IBM SPSS St	atistics									_	
Arquivo <u>E</u> ditar <u>V</u> isualizar <u>D</u> ados <u>T</u> ransforma	ar <u>I</u> nserir F <u>o</u>	rmatar <u>A</u> nalis:	ar <u>G</u> ráficos	<u>U</u> tilitário	os E <u>x</u> ten	nsões Jan	ela Ajuda				
😑 H 🖨 👌 🕗 💷 🗠	ີ 🛛 🧮										
а		5	268 271	102	233	229.898	501				
Log Regressão logística		6	260 261	169	233	231.831	493				
- E Título		7	262 258	926	238	241.074	500				
🕀 Observações		8	254 254	133	247	246.867	501				
Resumo de processamento do caso Codificação de variável dependente		9	202 192	149	298	307.851	500				
Codificação de variáveis categóricas		10	180 191	288	338	326.712	518				
Bloco 0: Bloco Inicial											
		Diservado	la de Class	ificação^a P Satisfied) 1.	Previsto Poi	rcentagem correta		A		5	
Testes de Omnibus do Modelo de Coeficiente:	Passo 1	Satisfied .00	2	259	389	85.3				D	
Resumo do modelo Can Teste de Hosmer e Lemeshow		1.0	0 1	708	644	27.4					
Tabela de contingência para teste de Hosmer		Porcentagem g	lobal			58.1					
→ Tabela de Classificação	a. O val	or de recorte é .5	00								
				Var	iáveis na	a equação		(C		5.00	
			в	9 E	he/M	df	Sid	Evp(B)	95% C.I. pa	Superior	
	Passo 1ª	eov/1)	- 029	0.59	459	1	109.	062	959	1.077	
	1 4350 1	309	033	005	405	1	525	1 003	.009	1.014	
		education	.003	.505	149 320	4	.020	1.003	.555	1.014	
		education(1)	128	092	1 939	4	164	1 1 37	949	1 361	
		education(2)	388	092	17.903	1	000	1 474	1 231	1.763	
		education(3)	477	092	26.995	1	.000	1.612	1 346	1.930	
		cuucation(3)	.+//	.032	20.333	1	.000	2.819	2 347	3 386	
		education(4)	1.036	093	122.884			2.013	6.071	3.000	
		education(4)	1.036	.093	7.965	1	005	523			
	a Variá	education(4) Constante	1.036 648) no passo 1:	.093 .230 sex.age.eg	122.889 7.965 ducation.	1	.005	.523			
4	a. Variá	education(4) Constante ivel(is) inserida(s	1.036 648) no passo 1:	.093 .230 sex, age, eo	7.965 ducation.	1	.005	.523			



Aula 11: Estimação de Modelos de Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Esta tabela apresenta a informação sobre o efeito das variáveis independentes sobre a variável dependent.
- A primeira coisa a olhar é se a relação das IV's com a DV é estatísticamente significativa, i.e.
 - 'Sig.' < 0.01 / 0.05 / 0.1
- Neste caso, observa-se que as variáveis 'sex' e 'age' não são significativas.

ta *Saío	da1 [Documento1] - V	'isualizador d	do IBM SPSS S	tatistic	s											-	
Arquivo	<u>E</u> ditar ⊻isualiza	ır <u>D</u> ados	Transform	ar <u>I</u>	nserir F <u>o</u>	rmatar <u>A</u>	nalisar <u>G</u>	ráficos	<u>U</u> tilitários	Exten	nsões Jar	nela Ajuda					
2	8 🖨 🗟		🛄 🗠	` ~	¥ 🧮				b I								
а						5	268	271.10	2	233	229.898	501					
Log Regress	ão logística					6	260	261.16	9	233	231.831	493					
i 🖆 Títul	0					7	262	258.92	5	238	241.074	500					
🔂 Obs	ervações					8	254	254.13	3	247	246.867	501					
-Liji Res	umo de processame	nto do caso				9	202	192.14	э	298	307.851	500					
-Cod	ificação de variavel de ificações de variáveis	spendente categóricas				10	180	191.28	3	338	326.712	518					
- E Bloc	o 0: Bloco Inicial Título Tabela de Classifica Variáveis na equação	ção	~~~			T	abela de	Classifi	cação ^a	ovieto							
E Bloc	variaveis nao preser o 1: Método = Enter	ites na equa	çao					e,	ric	501510							
- @	Título					Observed		.00	1.0	0 Poi	rcentagem correta						
- 10	Testes de Omnibus	do Modelo de	e Coeficiente:		Passo 1	Satisfied	00	2250	1 7	289	85.3	-					
	Resumo do modelo	omechow			1 4330 1	Gaustica	1.00	170	2 6	544	27.4	_					
- 6	Tabela de contingên	cia para teste	e de Hosmer			Porcentar	em alobal	1700			58.1	_					
	Tabela de Classifica	ção			a. O val	or de recort	a á 500				50.1	-					
·+ 🛍	Variáveis na equação	5		1	u. 0 vai	of de record	, c .500										
				l r					Variá	áveis na	a equação		A				
														95% C.I. pa	ara EXP(B)		
							В	1 1	3.E.	Wald	df	Sig.	Exp(B)	Inferior	Superior		
					Passo 1 ^a	sex(1)		039	.058	.459	1	.498	.962	.859	1.077		
						age		003	.005	.405	1	.525	1.003	.993	1.014		
						education	1			149.320	4	.000					
				7		education	n(1) .	128	.092	1.938	1	.164	1.137	.949	1.361		
						education	1(2) .	388	.092	17.903	1	.000	1.474	1.231	1.763		
						education	1(3) .	477	.092	26.995	1	.000	1.612	1.346	1.930		
						education	1(4) 1.	036	.093	122.889	1	.000	2.819	2.347	3.386		
						Constant	e	648	.230	7.965	1	.005	.523				
					a. Variá	vel(is) inser	ida(s) no pa	sso 1: sex	, age, edu	ucation.							
				Ľ													
4				2 M													



Aula 11: Estimação de Modelos de Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- De seguida, queremos ir ver os coeficientes do modelo:
 - Ser homem diminui o logaritmo da chance (log odds) de estar satisfeito com as condições de trabalho em 0.39...
 - Por cada ano adicional de vida, o logaritmo da chance (log odds) de estar satisfeito com as condições de trabalho aumenta em 0.03.
 - Comparado com ter apenas o ension básico (ISCED1), ter uma licenciatura aumenta o logaritmo da chance (log odds) de estar satisfeito com as condições de trabalho em 1.036.





Aula 11: Estimação de Modelos de Regressão Logística

- ISTO NÃO É NADA INTUITIVO!!
- Por isso é que nos interessa olhar para os odd-ratios (razão de chances)
- Interpretação:
 - Um odds ratio maior que 1 significa que um aumento em x leva a um aumento nas chances de y = 1.
 - Um odds ratio menor que 1 significa que um aumento em x leva a uma diminuição nas chances de y = 1.





Aula 11: Estimação de Modelos de Regressão Logística

Regressão Logística

 Melhor ainda, podemos calcular a mudança percentual nas odds por cada unidade adicional na IV:

Δ% das Odds = 100 *(Odds Ratio – 1)

- Exemplo:
 - As chances de estar satisfeito/a com as condições de trabalho são 3.8% - 100 * (0.962 - 1) - mais baixas para homens, em comparação com mulheres.
 - Cada aumento de um ano na idade leva a um aumento de 1,3% - 100 *(1,013-1) das chances de estar satisfeito/a com as condições de trabalho.

Optimized
Image:
$ \frac{5}{100} \frac{268}{271.102} \frac{233}{229.898} \frac{501}{501} \frac{5}{6} \frac{268}{260} \frac{271.102}{233} \frac{229.898}{231.831} \frac{493}{493} \frac{7}{2922} \frac{258.926}{238} \frac{231.831}{241.074} \frac{493}{500} \frac{7}{500} \frac{100}{200} \frac{8}{200} \frac{254}{254.133} \frac{224}{246.867} \frac{246.867}{501} \frac{500}{10} \frac{9}{10} \frac{202}{192.149} \frac{298}{307.851} \frac{307.851}{500} \frac{500}{10} \frac{100}{10} \frac{180}{191.288} \frac{326.712}{338} \frac{326.712}{518} \frac{518}{518} \frac{100}{10} \frac{100}{10} \frac{100}{100} \frac{100}{10} \frac{100}{$
ressão logística Titulo Observações Resumo de processamento do caso Codificação de variável dependente Codificação de variável dependente Codificação de variável dependente Codificação de variável dependente Codificação de variávels categóricas Bloco 0: Bloco Inicial → Tabela de Classificação → Variáveis na equação Bloco 1: Método = Enter → Tabela de Continuitus do Modelo de Coeficienter → Resumo do modelo → Tabela de Contingicia para teste de Hosmer → Tabela de Classificação → Variáveis na equação → Variáveis na equação
Titulo Observações Resumo de processamento do caso Codificação de variável dependente Codificação de variávels categóricas Bloco D: Bloco Inicial Tabela de Classificação Variáveis na equação Variáveis na equação Titulo Tabela de Classificação ^a Resumo de modelo Totale de contingéncia para teste de Hosmer Tabela de Classificação Nariáveis na equação Tabela de Classificação ^a Porcentagem Observado .00 1.00 1708 644 27.4 1.00 1708 644 27.4 1.00 1708 644 27.4 1.00 1708 644 27.4 1.00 58.1 a. O valor de recorte é .500
Observações Resumo de processamento do caso Codificações de variáveis categóricas Bloco 0: Bloco Inicial
Resumo de processamento do caso Codificação de variável dependente Codificação In Tabela de Classificação Variáveis não presentes na equação Previsto Bloco 1: Método = Enter In Tabela de Classificação de coeficiente: Resumo do modelo In Tabela de contingência para teste de Hosmer In Tabela de Classificação Nativeis na equação Passo 1 Satisfied Porcentagem Observado 0.00 1.00 1708 Fastes de Classificação Nativeis na equação In Tabela de classificação In Variáveis na equação In Variáveis na equação In Variáveis na equação
10 180 191.288 338 326.712 518 Bloco 0: Bloco Inicial Image: State
Observado .00 1.00 200 000 000 000 Image: Statistic and a statististic and a statis
- ☐ Titulo - ☐ Titulo - ☐ Titulo - ☐ Tabela de Classificação ^a - ☐ Variáveis na equação - ☐ Variáveis na equação - ☐ Tabela de Classificação ^a - ☐ Tabela de Classificação ^a - ☐ Tabela de Classificação ^a - ☐ Tabela de Consultation de Coeficiente: - ☐ Tabela de Classificação - ☐ Dabela de Classificação - ☐ Tabela de Classificação - ☐ Tabela de Classificação - ☐ Dabela de Classificação - ☐ Dabela de Classificação - ☐ Dab
Image: Tabela de Classificação Image: Tabela de Classificação <td< td=""></td<>
Image: Constraint of the constraint
Bloco 1: Método = Enter → Titulo → Testes de Omnibus do Modelo de Coeficiente: → Resumo do modelo → Tabela de Colassificação → Warláveis na equação B S.E. Wald df Sig. Exo(B) Inferior Superior
Image: Second
Passo 1 Satisfied 00 2259 389 85.3 Teste de Hosmer e Lemeshow Tabela de contingência para teste de Hosmer Tabela de classificação Tabela de Classificação Contrativeis na equação B S.E. Wald df Sig. Exo(B) Inferior Superior
Image: Constraint of the second se
Porcentagem global 58.1 Imabela de classificação a. O valor de recorte é .500 Image: Porcentagem global 58.1 Image: Porcentagem global 58.1 <t< td=""></t<>
a. O valor de recorte é .500 Variáveis na equação B S.E. Wald dr Sig. Exp(B) Inferior Superior
Variáveis na equação B S.E. Wald df Sig. Exp(B)
B S.E. Wald df Sig. Exp(B) Inferior Superior
B S.E. Wald df Sig. Exp(B) Inferior
Passo 1 ^a sex(1)039 .058 .459 1 .498 .962 .859 1.077
age .003 .005 .405 1 .525 1.003 .993 1.014
education 149.320 4 .000
→ education(1) .128 .092 1.938 1 .164 1.137 .949 1.361
education(2) .388 .092 17.903 1 .000 1.474 1.231 1.763
education(3) 477 092 26.995 1 000 1.612 1.346 1.930
education(4) 1.036 093 122.889 1 0.00 2.819 2.347 3.386
Constanta - 548 230 7 955 1 005 523
a. variaranja moendajo no passo r. son, ago, oducation.



Aula 11: Estimação de Modelos de Regressão Logística

Implementação do Modelo de Regressão Logística

6. Validar pressupostos do modelo de regressão logística



Aula 11: Estimação de Modelos de Regressão Logística

- Diagnósticos do modelo de regressão logística
 - Linearidade da relação entre os variáveis independentes e a variável dependente
 - Existência de Observações Influentes
 - Ausência de multicolinearidade



Aula 11: Estimação de Modelos de Regressão Logística

The End