

Regressão múltipla

Quando o passado permite prever o futuro, há um método estatístico denominado Regressão Linear que busca o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Em estatística, regressão é uma técnica que permite explorar e inferir a relação de uma variável dependente (variável de resposta) com variáveis independentes específicas (variáveis explicatórias). A análise da regressão pode ser usada como um método descritivo da análise de dados (como, por exemplo, o ajustamento de curvas) sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados. Regressão designa também uma equação matemática que descreva a relação entre duas ou mais variáveis

Quando há apenas 2 variáveis em análise diz-se "regressão simples" e quando há mais do que 2 variáveis envolvidas, diz-se "regressão múltipla". Na fórmula original tem-se

y = Xb + e

, que deve ser lido como o parametro matricial b é o que ajusta os y e os x observados.

A fórmula de cálculo de b é b = (X'X)^-1 X'y, onde X' é a matriz transposta de X e X^-1 é a matriz inversa de X.

Acompanhe o exemplo

Suponhamos querer estudar o comportamento do consumo de bens em uma determinada família, sujeita a dois parâmetros:

- renda e
- taxa de juros

Dos estudos de econometria, sabe-se que o consumo aumenta em função da renda e diminui em função da taxa de juros.

O trabalho começa pelo levantamento de dados prévios e obteve-se a seguinte tabela:

medida	consumo y	renda x1	tx juros x2
1	122	139	11,5%
2	114	126	12,0%
3	86	90	10,5%
4	134	144	9,0%
5	146	163	10,0%
6	107	136	12,0%
7	68	61	10,5%
8	117	62	8,0%
9	71	41	10,0%
10	98	120	11,5%

A matriz X do nosso problema é:

1	139	0.115
1	126	0.12
1	90	0.105
1	144	0.09
1	163	0.1
1	136	0.12
1	61	0.105
1	62	0.08
1	41	0.1
1	120	0.115

Note que a primeira coluna sempre é formada de unidades e as demais colunas são as variáveis em estudo (no caso: renda e taxa de juros).

Tomando a matriz transposta fica-se com

1	1	1	1	1	1	1	1	1	1
139	126	90	144	163	136	61	62	41	120
0.11	0.12	0.10	0.09	0.1	0.12	0.10	0.08	0.1	0.11

Multiplicando uma pela outra, tem-se

10	1082	1.05
1082	132744	115.4
1.05	115.4	0.1118

E a matriz inversa desta é

7.225436859	0.00095983853	-68.85039418
0.00095983853	0.00007350542	-0.084886905
-68.85039418	-0.08488690598	743.1919753

Agora, multiplicando a inversa por X' fica:

-5	-.91	.08	1.16	.49	-.90	.05	1.7	.3	-.5
.0	.00	-.00	.00	.00	.0	-.00	-.0	-.0	.0
4.8	9.63	1.54	-14.1	-8.36	8.7	4.0	-14.6	1.9	6.4

Agora multiplicando esta matriz pela matriz coluna de y (consumo) que é (devidamente horizontalizada para economizar papel)

122	114	86	134	146	107	68	117	71	98
-----	-----	----	-----	-----	-----	----	-----	----	----

Fica

148.5220194
0.6135933813
-1034.40784

Interpretação

Neste problema, o consumo é

Consumo =

148,52 + 0.6135 x Renda - 1.034.41 x TxJuros

Ou seja, tirando a parte do Consumo que não é influenciada pela Taxa de Juros, o incremento de \$ 1 na Renda causa um incremento esperado de \$ 0,6136 no Consumo; além disso, o incremento de 1 ponto percentual (0,01) na Taxa de Juros causa um decréscimo esperado de \$ 10,3441 no Consumo.

Para você fazer

1. Considere o mesmo problema com os seguintes valores obtidos

medida	consumo y	renda x1	tx juros x2
1	95	115	.1110
2	91	114	.0930
3	86	112	.1000
4	98	127	.1180
5	84	111	.1160
6	86	112	.1060
7	90	103	.1000
8	97	101	.1140
9	109	110	.0950
10	95	122	.1180

Ache os parâmetros de renda e taxa de juros para o problema proposto.

t. fixo	renda	tx de juros
---------	-------	-------------

2. Mais um problema, agora envolvendo um outro modelo econométrico (obviamente FICTICIO) no qual estão relacionadas 3 variáveis independentes e uma dependente. Não se esqueça que a matemática lava as mãos sobre essas definições. É coisa do planejador. São elas:

t	dep.	consumo de whisky importado (em caixas)
x	indep.	venda de propriedades em Caiobá (em um)
y	indep.	número de falências (sentido oposto, em um)
z	indep.	número de BMW licenciadas (em um)

O objetivo então é achar

t = a + b.x + c.y + d.z

No último ano foram levantados os valores de x, y e z e obteve-se

medida	t	x	y	z
1	1950696	637	571	767
2	3709571	835	35	624
3	2811277	303	162	735
4	186519	724	660	371
5	-947222	880	931	277
6	-2934127	195	841	-28
7	989749	796	329	232
8	-411222	171	222	34
9	-1649218	777	747	-31
10	2487374	741	409	707

Usando a teoria acima, você deve achar os valores de a, b, c e d.

a	b	c	d
---	---	---	---

3. O problema agora é que você é o dono do maior restaurante de Santa Felicidade. Precisa prever a quantidade de refeições a serem servidas no próximo Dia das Mães. Depois de muito estudar, você contratou um engenheiro que levantou dados dos últimos 6 anos e elaborou o seguinte modelo econométrico:

a	indep	número de refeições a serem servidas no próximo dia das mães
b	dep	número de casamentos em Curitiba no ano anterior
c	dep	renda média das famílias no Paraná no ano anterior, em reais
d	dep	frangos abatidos na última semana do mês de abril no Paraná
e	dep	veículos emplacadas na ciretran de S. Felicidade no ano anterior

A fórmula do modelo é:

a = α + βb + γc + δd + ε

Os dados:

	a	b	c	d	e
1	7632	12898	7328	153716	6421
2	8512	13760	7818	169920	8621
3	7365	14217	10502	140232	8365
4	7933	17247	8350	160194	4543
5	7553	17534	13890	147259	3695
6	8235	19207	11743	153206	10293

Achar os valores de α, β, γ, δ e ε.

α	β	γ	δ	ε
---	---	---	---	---

Observação importante: No Python, ao gerar as matrizes, não esquecer de pedir campos float, porque senão ele pode resolver usando inteiros e vai achar respostas erradas...



122-68993 - f/na

Regressão múltipla

Quando o passado permite prever o futuro, há um método estatístico denominado Regressão Linear que busca o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Em estatística, regressão é uma técnica que permite explorar e inferir a relação de uma variável dependente (variável de resposta) com variáveis independentes específicas (variáveis explicatórias). A análise da regressão pode ser usada como um método descritivo da análise de dados (como, por exemplo, o ajustamento de curvas) sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados. Regressão designa também uma equação matemática que descreva a relação entre duas ou mais variáveis

Quando há apenas 2 variáveis em análise diz-se "regressão simples" e quando há mais do que 2 variáveis envolvidas, diz-se "regressão múltipla". Na fórmula original tem-se

y = Xb + ε

, que deve ser lido como o parametro matricial b é o que ajusta os y e os x observados.

A fórmula de cálculo de b é b = (X'X)⁻¹X'y, onde X' é a matriz transposta de X e X⁻¹ é a matriz inversa de X.

Acompanhe o exemplo

Suponhamos querer estudar o comportamento do consumo de bens em uma determinada família, sujeita a dois parâmetros:

- renda e
- taxa de juros

Dos estudos de econometria, sabe-se que o consumo aumenta em função da renda e diminui em função da taxa de juros.

O trabalho começa pelo levantamento de dados prévios e obteve-se a seguinte tabela:

medida	consumo y	renda x ₁	tx juros x ₂
1	122	139	11,5%
2	114	126	12,0%
3	86	90	10,5%
4	134	144	9,0%
5	146	163	10,0%
6	107	136	12,0%
7	68	61	10,5%
8	117	62	8,0%
9	71	41	10,0%
10	98	120	11,5%

A matriz X do nosso problema é:

1	139	0.115
1	126	0.12
1	90	0.105
1	144	0.09
1	163	0.1
1	136	0.12
1	61	0.105
1	62	0.08
1	41	0.1
1	120	0.115

Note que a primeira coluna sempre é formada de unidades e as demais colunas são as variáveis em estudo (no caso: renda e taxa de juros).

Tomando a matriz transposta fica-se com

1	1	1	1	1	1	1	1	1	1
139	126	90	144	163	136	61	62	41	120
0.11	0.12	0.10	0.09	0.1	0.12	0.10	0.08	0.1	0.11

Multiplicando uma pela outra, tem-se

10	1082	1.05
1082	132744	115.4
1.05	115.4	0.1118

E a matriz inversa desta é

7.225436859	0.00095983853	-68.85039418
0.00095983853	0.0007350542	-0.084886905
-68.85039418	-0.08488690598	743.1919753

Agora, multiplicando a inversa por X' fica:

-5	-.91	.08	1.16	.49	-.90	.05	1.7	.3	-.5
.0	.00	-.00	.00	.00	.0	-.00	-.0	-.0	.0
4.8	9.63	1.54	-14.1	-8.36	8.7	4.0	-14.6	1.9	6.4

Agora multiplicando esta matriz pela matriz coluna de y (consumo) que é (devidamente horizontalizada para economizar papel)

122	114	86	134	146	107	68	117	71	98
-----	-----	----	-----	-----	-----	----	-----	----	----

Fica

148.5220194
0.6135933813
-1034.40784

Interpretação

Neste problema, o consumo é

Consumo =

148,52 + 0.6135 × Renda - 1.034.41 × TxJuros

Ou seja, tirando a parte do Consumo que não é influenciada pela Taxa de Juros, o incremento de \$ 1 na Renda causa um incremento esperado de \$ 0,6136 no Consumo; além disso, o incremento de 1 ponto percentual (0,01) na Taxa de Juros causa um decréscimo esperado de \$ 10,3441 no Consumo.

Para você fazer

1. Considere o mesmo problema com os seguintes valores obtidos

medida	consumo y	renda x ₁	tx juros x ₂
1	93	130	.1080
2	94	117	.1000
3	88	125	.1140
4	105	123	.1160
5	87	127	.0980
6	88	114	.0960
7	98	129	.1110
8	93	111	.1050
9	96	128	.1030
10	92	117	.1120

Ache os parâmetros de renda e taxa de juros para o problema proposto.

t. fixo	renda	tx de juros
---------	-------	-------------

2. Mais um problema, agora envolvendo um outro modelo econométrico (obviamente FICTICIO) no qual estão relacionadas 3 variáveis independentes e uma dependente. Não se esqueça que a matemática lava as mãos sobre essas definições. É coisa do planejador. São elas:

t	dep.	consumo de whisky importado (em caixas)
x	indep.	venda de propriedades em Caiobá (em um)
y	indep.	número de falências (sentido oposto, em um)
z	indep.	número de BMW licenciadas (em um)

O objetivo então é achar

t = a + b.x + c.y + d.z

No último ano foram levantados os valores de x, y e z e obteve-se

medida	t	x	y	z
1	-421612	391	502	13
2	2537085	-78	274	362
3	-1234329	544	701	-61
4	2462645	394	847	427
5	3671665	262	773	570
6	-316231	350	520	31
7	4459678	445	462	606
8	1123700	95	586	232
9	45716	855	576	61
10	6279957	-46	355	836

Usando a teoria acima, você deve achar os valores de a, b, c e d.

a	b	c	d
---	---	---	---

3. O problema agora é que você é o dono do maior restaurante de Santa Felicidade. Precisa prever a quantidade de refeições a serem servidas no próximo Dia das Mães. Depois de muito estudar, você contratou um engenheiro que levantou dados dos últimos 6 anos e elaborou o seguinte modelo econométrico:

a	indep	número de refeições a serem servidas no próximo dia das mães
b	dep	número de casamentos em Curitiba no ano anterior
c	dep	renda média das famílias no Paraná no ano anterior, em reais
d	dep	frangos abatidos na última semana do mês de abril no Paraná
e	dep	veículos emplacadas na ciretran de S. Felicidade no ano anterior

A fórmula do modelo é:

a = α + βb + γc + δd + ε

Os dados:

	a	b	c	d	e
1	6045	11769	12728	141072	9181
2	5058	12310	12806	116600	7957
3	6323	13448	11031	154776	4295
4	6688	15925	10604	156712	10043
5	5562	16716	10561	132077	6094
6	3950	17466	10987	83903	10433

Achar os valores de α, β, γ, δ e ε.

α	β	γ	δ	ε
---	---	---	---	---

Observação importante: No Python, ao gerar as matrizes, não esquecer de pedir campos float, porque senão ele pode resolver usando inteiros e vai achar respostas erradas...



122-69789 - f/na

Regressão múltipla

Quando o passado permite prever o futuro, há um método estatístico denominado Regressão Linear que busca o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Em estatística, regressão é uma técnica que permite explorar e inferir a relação de uma variável dependente (variável de resposta) com variáveis independentes específicas (variáveis explicatórias). A análise da regressão pode ser usada como um método descritivo da análise de dados (como, por exemplo, o ajustamento de curvas) sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados. Regressão designa também uma equação matemática que descreva a relação entre duas ou mais variáveis

Quando há apenas 2 variáveis em análise diz-se "regressão simples" e quando há mais do que 2 variáveis envolvidas, diz-se "regressão múltipla". Na fórmula original tem-se

y = Xb + ε

, que deve ser lido como o parametro matricial b é o que ajusta os y e os x observados.

A fórmula de cálculo de b é b = (X'X)⁻¹X'y, onde X' é a matriz transposta de X é X⁻¹ é a matriz inversa de X.

Acompanhe o exemplo

Suponhamos querer estudar o comportamento do consumo de bens em uma determinada família, sujeita a dois parâmetros:

- renda e
- taxa de juros

Dos estudos de econometria, sabe-se que o consumo aumenta em função da renda e diminui em função da taxa de juros.

O trabalho começa pelo levantamento de dados prévios e obteve-se a seguinte tabela:

medida	consumo y	renda x ₁	tx juros x ₂
1	122	139	11,5%
2	114	126	12,0%
3	86	90	10,5%
4	134	144	9,0%
5	146	163	10,0%
6	107	136	12,0%
7	68	61	10,5%
8	117	62	8,0%
9	71	41	10,0%
10	98	120	11,5%

A matriz X do nosso problema é:

1	139	0.115
1	126	0.12
1	90	0.105
1	144	0.09
1	163	0.1
1	136	0.12
1	61	0.105
1	62	0.08
1	41	0.1
1	120	0.115

Note que a primeira coluna sempre é formada de unidades e as demais colunas são as variáveis em estudo (no caso: renda e taxa de juros).

Tomando a matriz transposta fica-se com

1	1	1	1	1	1	1	1	1	1
139	126	90	144	163	136	61	62	41	120
0.11	0.12	0.10	0.09	0.1	0.12	0.10	0.08	0.1	0.11

Multiplicando uma pela outra, tem-se

10	1082	1.05
1082	132744	115.4
1.05	115.4	0.1118

E a matriz inversa desta é

7.225436859	0.00095983853	-68.85039418
0.00095983853	0.0007350542	-0.084886905
-68.85039418	-0.08488690598	743.1919753

Agora, multiplicando a inversa por X' fica:

-5	-.91	.08	1.16	.49	-.90	.05	1.7	.3	-.5
.0	.00	-.00	.00	.00	.0	-.00	-.0	-.0	.0
4.8	9.63	1.54	-14.1	-8.36	8.7	4.0	-14.6	1.9	6.4

Agora multiplicando esta matriz pela matriz coluna de y (consumo) que é (devidamente horizontalizada para economizar papel)

122	114	86	134	146	107	68	117	71	98
-----	-----	----	-----	-----	-----	----	-----	----	----

Fica

148.5220194
0.6135933813
-1034.40784

Interpretação

Neste problema, o consumo é

Consumo =

148,52 + 0.6135 × Renda - 1.034.41 × TxJuros

Ou seja, tirando a parte do Consumo que não é influenciada pela Taxa de Juros, o incremento de \$ 1 na Renda causa um incremento esperado de \$ 0,6136 no Consumo; além disso, o incremento de 1 ponto percentual (0,01) na Taxa de Juros causa um decréscimo esperado de \$ 10,3441 no Consumo.

Para você fazer

1. Considere o mesmo problema com os seguintes valores obtidos

medida	consumo y	renda x ₁	tx juros x ₂
1	98	115	.1150
2	104	129	.1060
3	83	121	.1190
4	96	119	.1110
5	92	129	.0920
6	105	104	.1170
7	98	117	.0950
8	104	111	.1190
9	100	103	.0930
10	85	107	.1140

Ache os parâmetros de renda e taxa de juros para o problema proposto.

t. fixo	renda	tx de juros
---------	-------	-------------

2. Mais um problema, agora envolvendo um outro modelo econométrico (obviamente FICTICIO) no qual estão relacionadas 3 variáveis independentes e uma dependente. Não se esqueça que a matemática lava as mãos sobre essas definições. É coisa do planejador. São elas:

t	dep.	consumo de whisky importado (em caixas)
x	indep.	venda de propriedades em Caiobá (em um)
y	indep.	número de falências (sentido oposto, em um)
z	indep.	número de BMW licenciadas (em um)

O objetivo então é achar

t = a + b.x + c.y + d.z

No último ano foram levantados os valores de x, y e z e obteve-se

medida	t	x	y	z
1	1093694	408	447	632
2	-31553	899	844	20
3	-161550	383	723	209
4	2009166	842	95	544
5	475070	487	799	527
6	-909067	381	930	-19
7	-272812	178	893	420
8	230011	427	784	429
9	669456	53	609	776
10	1058024	657	9	110

Usando a teoria acima, você deve achar os valores de a, b, c e d.

a	b	c	d
---	---	---	---

3. O problema agora é que você é o dono do maior restaurante de Santa Felicidade. Precisa prever a quantidade de refeições a serem servidas no próximo Dia das Mães. Depois de muito estudar, você contratou um engenheiro que levantou dados dos últimos 6 anos e elaborou o seguinte modelo econométrico:

a	indep	número de refeições a serem servidas no próximo dia das mães
b	dep	número de casamentos em Curitiba no ano anterior
c	dep	renda média das famílias no Paraná no ano anterior, em reais
d	dep	frangos abatidos na última semana do mês de abril no Paraná
e	dep	veículos emplacadas na ciretran de S. Felicidade no ano anterior

A fórmula do modelo é:

a = α + βb + γc + δd + ε

Os dados:

	a	b	c	d	e
1	6757	14297	7158	103191	4648
2	9250	15063	7455	142050	8196
3	8883	16961	11186	128430	10555
4	6682	17283	9819	90986	10041
5	7181	17287	8367	110829	3512
6	10613	19676	13494	158261	9581

Achar os valores de α, β, γ, δ e ε.

α	β	γ	δ	ε
---	---	---	---	---

Observação importante: No Python, ao gerar as matrizes, não esquecer de pedir campos float, porque senão ele pode resolver usando inteiros e vai achar respostas erradas...



122-69008 - f/na

Regressão múltipla

Quando o passado permite prever o futuro, há um método estatístico denominado Regressão Linear que busca o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Em estatística, regressão é uma técnica que permite explorar e inferir a relação de uma variável dependente (variável de resposta) com variáveis independentes específicas (variáveis explicatórias). A análise da regressão pode ser usada como um método descritivo da análise de dados (como, por exemplo, o ajustamento de curvas) sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados. Regressão designa também uma equação matemática que descreva a relação entre duas ou mais variáveis

Quando há apenas 2 variáveis em análise diz-se "regressão simples" e quando há mais do que 2 variáveis envolvidas, diz-se "regressão múltipla". Na fórmula original tem-se

y = Xb + e

, que deve ser lido como o parametro matricial b é o que ajusta os y e os x observados.

A fórmula de cálculo de b é b = (X'X)^-1 X'y, onde X' é a matriz transposta de X e X^-1 é a matriz inversa de X.

Acompanhe o exemplo

Suponhamos querer estudar o comportamento do consumo de bens em uma determinada família, sujeita a dois parâmetros:

- renda e
- taxa de juros

Dos estudos de econometria, sabe-se que o consumo aumenta em função da renda e diminui em função da taxa de juros.

O trabalho começa pelo levantamento de dados prévios e obteve-se a seguinte tabela:

medida	consumo y	renda x1	tx juros x2
1	122	139	11,5%
2	114	126	12,0%
3	86	90	10,5%
4	134	144	9,0%
5	146	163	10,0%
6	107	136	12,0%
7	68	61	10,5%
8	117	62	8,0%
9	71	41	10,0%
10	98	120	11,5%

A matriz X do nosso problema é:

1	139	0.115
1	126	0.12
1	90	0.105
1	144	0.09
1	163	0.1
1	136	0.12
1	61	0.105
1	62	0.08
1	41	0.1
1	120	0.115

Note que a primeira coluna sempre é formada de unidades e as demais colunas são as variáveis em estudo (no caso: renda e taxa de juros).

Tomando a matriz transposta fica-se com

1	1	1	1	1	1	1	1	1	1
139	126	90	144	163	136	61	62	41	120
0.11	0.12	0.10	0.09	0.1	0.12	0.10	0.08	0.1	0.11

Multiplicando uma pela outra, tem-se

10	1082	1.05
1082	132744	115.4
1.05	115.4	0.1118

E a matriz inversa desta é

7.225436859	0.00095983853	-68.85039418
0.00095983853	0.00007350542	-0.084886905
-68.85039418	-0.08488690598	743.1919753

Agora, multiplicando a inversa por X' fica:

-5	-.91	.08	1.16	.49	-.90	.05	1.7	.3	-.5
.0	.00	-.00	.00	.00	.0	-.00	-.0	-.0	.0
4.8	9.63	1.54	-14.1	-8.36	8.7	4.0	-14.6	1.9	6.4

Agora multiplicando esta matriz pela matriz coluna de y (consumo) que é (devidamente horizontalizada para economizar papel)

122	114	86	134	146	107	68	117	71	98
-----	-----	----	-----	-----	-----	----	-----	----	----

Fica

148.5220194
0.6135933813
-1034.40784

Interpretação

Neste problema, o consumo é

Consumo =

148,52 + 0.6135 x Renda - 1.034.41 x TxJuros

Ou seja, tirando a parte do Consumo que não é influenciada pela Taxa de Juros, o incremento de \$ 1 na Renda causa um incremento esperado de \$ 0,6136 no Consumo; além disso, o incremento de 1 ponto percentual (0,01) na Taxa de Juros causa um decréscimo esperado de \$ 10,3441 no Consumo.

Para você fazer

1. Considere o mesmo problema com os seguintes valores obtidos

medida	consumo y	renda x1	tx juros x2
1	92	126	.0940
2	85	118	.0990
3	99	122	.0910
4	102	122	.0970
5	87	130	.1080
6	90	111	.0910
7	86	118	.1170
8	108	116	.1190
9	109	130	.1170
10	94	105	.0990

Ache os parâmetros de renda e taxa de juros para o problema proposto.

t. fixo	renda	tx de juros
---------	-------	-------------

2. Mais um problema, agora envolvendo um outro modelo econométrico (obviamente FICTICIO) no qual estão relacionadas 3 variáveis independentes e uma dependente. Não se esqueça que a matemática lava as mãos sobre essas definições. É coisa do planejador. São elas:

t	dep.	consumo de whisky importado (em caixas)
x	indep.	venda de propriedades em Caiobá (em um)
y	indep.	número de falências (sentido oposto, em um)
z	indep.	número de BMW licenciadas (em um)

O objetivo então é achar

t = a + b.x + c.y + d.z

No último ano foram levantados os valores de x, y e z e obteve-se

medida	t	x	y	z
1	751152	112	198	218
2	-384000	434	855	381
3	4148582	470	263	686
4	279550	832	624	193
5	-1694564	90	669	159
6	-102553	39	7	-26
7	1659831	815	546	366
8	471567	171	202	156
9	1237442	168	62	183
10	-658452	808	423	-88

Usando a teoria acima, você deve achar os valores de a, b, c e d.

a	b	c	d
---	---	---	---

3. O problema agora é que você é o dono do maior restaurante de Santa Felicidade. Precisa prever a quantidade de refeições a serem servidas no próximo Dia das Mães. Depois de muito estudar, você contratou um engenheiro que levantou dados dos últimos 6 anos e elaborou o seguinte modelo econométrico:

a	indep	número de refeições a serem servidas no próximo dia das mães
b	dep	número de casamentos em Curitiba no ano anterior
c	dep	renda média das famílias no Paraná no ano anterior, em reais
d	dep	frangos abatidos na última semana do mês de abril no Paraná
e	dep	veículos emplacadas na ciretran de S. Felicidade no ano anterior

A fórmula do modelo é:

a = α + βb + γc + δd + ε

Os dados:

	a	b	c	d	e
1	10212	11850	8459	129420	10735
2	9516	15584	12846	117795	6164
3	11761	16693	10451	151806	8021
4	12265	17455	13537	158549	6055
5	12249	19665	11444	162288	3105
6	8856	19808	13952	102082	8138

Achar os valores de α, β, γ, δ e ε.

α	β	γ	δ	ε
---	---	---	---	---

Observação importante: No Python, ao gerar as matrizes, não esquecer de pedir campos float, porque senão ele pode resolver usando inteiros e vai achar respostas erradas...



122-69015 - f/na

Regressão múltipla

Quando o passado permite prever o futuro, há um método estatístico denominado Regressão Linear que busca o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Em estatística, regressão é uma técnica que permite explorar e inferir a relação de uma variável dependente (variável de resposta) com variáveis independentes específicas (variáveis explicatórias). A análise da regressão pode ser usada como um método descritivo da análise de dados (como, por exemplo, o ajustamento de curvas) sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados. Regressão designa também uma equação matemática que descreva a relação entre duas ou mais variáveis

Quando há apenas 2 variáveis em análise diz-se "regressão simples" e quando há mais do que 2 variáveis envolvidas, diz-se "regressão múltipla". Na fórmula original tem-se

y = Xb + e

, que deve ser lido como o parametro matricial b é o que ajusta os y e os x observados.

A fórmula de cálculo de b é b = (X'X)^-1 X'y, onde X' é a matriz transposta de X e X^-1 é a matriz inversa de X.

Acompanhe o exemplo

Suponhamos querer estudar o comportamento do consumo de bens em uma determinada família, sujeita a dois parâmetros:

- renda e
- taxa de juros

Dos estudos de econometria, sabe-se que o consumo aumenta em função da renda e diminui em função da taxa de juros.

O trabalho começa pelo levantamento de dados prévios e obteve-se a seguinte tabela:

medida	consumo y	renda x1	tx juros x2
1	122	139	11,5%
2	114	126	12,0%
3	86	90	10,5%
4	134	144	9,0%
5	146	163	10,0%
6	107	136	12,0%
7	68	61	10,5%
8	117	62	8,0%
9	71	41	10,0%
10	98	120	11,5%

A matriz X do nosso problema é:

1	139	0.115
1	126	0.12
1	90	0.105
1	144	0.09
1	163	0.1
1	136	0.12
1	61	0.105
1	62	0.08
1	41	0.1
1	120	0.115

Note que a primeira coluna sempre é formada de unidades e as demais colunas são as variáveis em estudo (no caso: renda e taxa de juros).

Tomando a matriz transposta fica-se com

1	1	1	1	1	1	1	1	1	1
139	126	90	144	163	136	61	62	41	120
0.11	0.12	0.10	0.09	0.1	0.12	0.10	0.08	0.1	0.11

Multiplicando uma pela outra, tem-se

10	1082	1.05
1082	132744	115.4
1.05	115.4	0.1118

E a matriz inversa desta é

7.225436859	0.00095983853	-68.85039418
0.00095983853	0.0007350542	-0.084886905
-68.85039418	-0.08488690598	743.1919753

Agora, multiplicando a inversa por X' fica:

-5	-.91	.08	1.16	.49	-.90	.05	1.7	.3	-.5
.0	.00	-.00	.00	.00	.0	-.00	-.0	-.0	.0
4.8	9.63	1.54	-14.1	-8.36	8.7	4.0	-14.6	1.9	6.4

Agora multiplicando esta matriz pela matriz coluna de y (consumo) que é (devidamente horizontalizada para economizar papel)

122	114	86	134	146	107	68	117	71	98
-----	-----	----	-----	-----	-----	----	-----	----	----

Fica

148.5220194
0.6135933813
-1034.40784

Interpretação

Neste problema, o consumo é

Consumo =

148,52 + 0.6135 x Renda - 1.034.41 x TxJuros

Ou seja, tirando a parte do Consumo que não é influenciada pela Taxa de Juros, o incremento de \$ 1 na Renda causa um incremento esperado de \$ 0,6136 no Consumo; além disso, o incremento de 1 ponto percentual (0,01) na Taxa de Juros causa um decréscimo esperado de \$ 10,3441 no Consumo.

Para você fazer

1. Considere o mesmo problema com os seguintes valores obtidos

medida	consumo y	renda x1	tx juros x2
1	100	130	.1180
2	110	110	.1100
3	95	115	.1150
4	83	101	.1160
5	100	127	.1130
6	94	106	.0930
7	107	125	.0910
8	88	102	.1010
9	100	128	.0990
10	82	119	.1200

Ache os parâmetros de renda e taxa de juros para o problema proposto.

t. fixo	renda	tx de juros
---------	-------	-------------

2. Mais um problema, agora envolvendo um outro modelo econométrico (obviamente FICTICIO) no qual estão relacionadas 3 variáveis independentes e uma dependente. Não se esqueça que a matemática lava as mãos sobre essas definições. É coisa do planejador. São elas:

t	dep.	consumo de whisky importado (em caixas)
x	indep.	venda de propriedades em Caiobá (em um)
y	indep.	número de falências (sentido oposto, em um)
z	indep.	número de BMW licenciadas (em um)

O objetivo então é achar

t = a + b.x + c.y + d.z

No último ano foram levantados os valores de x, y e z e obteve-se

medida	t	x	y	z
1	736180	-9	721	858
2	-3150792	177	880	160
3	-2912874	286	771	-1
4	-299201	167	656	455
5	4935115	414	361	882
6	-1714161	352	823	208
7	7753552	898	76	681
8	3112949	305	185	475
9	191586	173	312	194
10	1166122	174	660	715

Usando a teoria acima, você deve achar os valores de a, b, c e d.

a	b	c	d
---	---	---	---

3. O problema agora é que você é o dono do maior restaurante de Santa Felicidade. Precisa prever a quantidade de refeições a serem servidas no próximo Dia das Mães. Depois de muito estudar, você contratou um engenheiro que levantou dados dos últimos 6 anos e elaborou o seguinte modelo econométrico:

a	indep	número de refeições a serem servidas no próximo dia das mães
b	dep	número de casamentos em Curitiba no ano anterior
c	dep	renda média das famílias no Paraná no ano anterior, em reais
d	dep	frangos abatidos na última semana do mês de abril no Paraná
e	dep	veículos emplacadas na ciretran de S. Felicidade no ano anterior

A fórmula do modelo é:

a = α + βb + γc + δd + ε

Os dados:

	a	b	c	d	e
1	11208	11821	8958	167762	6865
2	8414	13279	10973	118938	6085
3	8614	14610	11212	118318	8924
4	9082	16641	8190	128323	8051
5	11362	16652	11825	168256	4358
6	6462	17164	11368	81082	8200

Achar os valores de α, β, γ, δ e ε.

α	β	γ	δ	ε
---	---	---	---	---

Observação importante: No Python, ao gerar as matrizes, não esquecer de pedir campos float, porque senão ele pode resolver usando inteiros e vai achar respostas erradas...



122-69022 - f/na

Regressão múltipla

Quando o passado permite prever o futuro, há um método estatístico denominado Regressão Linear que busca o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Em estatística, regressão é uma técnica que permite explorar e inferir a relação de uma variável dependente (variável de resposta) com variáveis independentes específicas (variáveis explicatórias). A análise da regressão pode ser usada como um método descritivo da análise de dados (como, por exemplo, o ajustamento de curvas) sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados. Regressão designa também uma equação matemática que descreva a relação entre duas ou mais variáveis

Quando há apenas 2 variáveis em análise diz-se "regressão simples" e quando há mais do que 2 variáveis envolvidas, diz-se "regressão múltipla".

Na fórmula original tem-se

$$y = Xb + \epsilon$$

, que deve ser lido como o parametro matricial b é o que ajusta os y e os x observados.

A fórmula de cálculo de b é $b = (X'X)^{-1}X'y$, onde X' é a matriz transposta de X e X^{-1} é a matriz inversa de X .

Acompanhe o exemplo

Suponhamos querer estudar o comportamento do consumo de bens em uma determinada família, sujeita a dois parâmetros:

- renda e
- taxa de juros

Dos estudos de econometria, sabe-se que o consumo aumenta em função da renda e diminui em função da taxa de juros.

O trabalho começa pelo levantamento de dados prévios e obteve-se a seguinte tabela:

medida	consumo y	renda x_1	tx juros x_2
1	122	139	11,5%
2	114	126	12,0%
3	86	90	10,5%
4	134	144	9,0%
5	146	163	10,0%
6	107	136	12,0%
7	68	61	10,5%
8	117	62	8,0%
9	71	41	10,0%
10	98	120	11,5%

A matriz X do nosso problema é:

```
1 139 0.115
1 126 0.12
1 90 0.105
1 144 0.09
1 163 0.1
1 136 0.12
1 61 0.105
1 62 0.08
1 41 0.1
1 120 0.115
```

Note que a primeira coluna sempre é formada de unidades e as demais colunas são as variáveis em estudo (no caso: renda e taxa de juros).

Tomando a matriz transposta fica-se com

```
1 1 1 1 1 1 1 1 1 1
139 126 90 144 163 136 61 62 41 120
0.11 0.12 0.10 0.09 0.1 0.12 0.10 0.08 0.1 0.11
```

Multiplicando uma pela outra, tem-se

```
10 1082 1.05
1082 132744 115.4
1.05 115.4 0.1118
```

E a matriz inversa desta é

```
7.225436859 0.00095983853 -68.85039418
0.00095983853 0.0007350542 -0.084886905
-68.85039418 -0.08488690598 743.1919753
```

Agora, multiplicando a inversa por X' fica:

```
-5 -91 .08 1.16 .49 -.90 .05 1.7 .3 -.5
.0 .00 -.00 .00 .00 .0 -.00 -.0 -.0 .0
4.8 9.63 1.54 -14.1 -8.36 8.7 4.0 -14.6 1.9 6.4
```

Agora multiplicando esta matriz pela matriz coluna de y (consumo) que é (devidamente horizontalizada para economizar papel)

```
122 114 86 134 146 107 68 117 71 98
Fica
```

```
148.5220194
0.6135933813
-1034.40784
```

Interpretação

Neste problema, o consumo é

$$Consumo =$$

$$148,52 + 0.6135 \times Renda - 1.034.41 \times TxJuros$$

Ou seja, tirando a parte do Consumo que não é influenciada pela Taxa de Juros, o incremento de \$ 1 na Renda causa um incremento esperado de \$ 0,6136 no Consumo; além disso, o incremento de 1 ponto percentual (0,01) na Taxa de Juros causa um decréscimo esperado de \$ 10,3441 no Consumo.

Para você fazer

1. Considere o mesmo problema com os seguintes valores obtidos

medida	consumo y	renda x_1	tx juros x_2
1	91	116	.1200
2	106	108	.1170
3	99	119	.1100
4	99	102	.1120
5	100	106	.1040
6	108	109	.1160
7	83	122	.1130
8	100	124	.1200
9	93	112	.0990
10	88	118	.1200

Ache os parâmetros de renda e taxa de juros para o problema proposto.

t. fixo	renda	tx de juros

2. Mais um problema, agora envolvendo um outro modelo econométrico (obviamente FICTICIO) no qual estão relacionadas 3 variáveis independentes e uma dependente. Não se esqueça que a matemática lava as mãos sobre essas definições. É coisa do planejador. São elas:

t	dep.	consumo de whisky importado (em caixas)
x	indep.	venda de propriedades em Caiobá (em um)
y	indep.	número de falências (sentido oposto, em um)
z	indep.	número de BMW licenciadas (em um)

O objetivo então é achar

$$t = a + b.x + c.y + d.z$$

No último ano foram levantados os valores de x , y e z e obteve-se

medida	t	x	y	z
1	2131647	68	817	740
2	1791407	408	828	553
3	-246146	809	943	78
4	-21079	90	579	226
5	140932	870	505	-75
6	2484521	894	966	559
7	2503212	543	88	293
8	2097031	777	885	495
9	2850149	580	376	475
10	1540622	110	126	298

Usando a teoria acima, você deve achar os valores de a , b , c e d .

a	b	c	d

3. O problema agora é que você é o dono do maior restaurante de Santa Felicidade. Precisa prever a quantidade de refeições a serem servidas no próximo Dia das Mães. Depois de muito estudar, você contratou um engenheiro que levantou dados dos últimos 6 anos e elaborou o seguinte modelo econométrico:

a	indep	número de refeições a serem servidas no próximo dia das mães
b	dep	número de casamentos em Curitiba no ano anterior
c	dep	renda média das famílias no Paraná no ano anterior, em reais
d	dep	frangos abatidos na última semana do mês de abril no Paraná
e	dep	veículos emplacadas na ciretran de S. Felicidade no ano anterior

A fórmula do modelo é:

$$a = \alpha + \beta b + \gamma c + \delta d + \epsilon$$

Os dados:

	a	b	c	d	e
1	5605	11210	13375	94736	4399
2	6450	12597	13536	103866	9218
3	8652	12727	6856	155544	9775
4	6288	14420	10437	104276	8235
5	5758	16191	10853	89183	9933
6	8148	19179	9269	150563	4107

Achar os valores de α , β , γ , δ e ϵ .

α	β	γ	δ	ϵ

Observação importante: No Python, ao gerar as matrizes, não esquecer de pedir campos float, porque senão ele pode resolver usando inteiros e vai achar respostas erradas...



122-69208 - f/na

Regressão múltipla

Quando o passado permite prever o futuro, há um método estatístico denominado Regressão Linear que busca o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Em estatística, regressão é uma técnica que permite explorar e inferir a relação de uma variável dependente (variável de resposta) com variáveis independentes específicas (variáveis explicatórias). A análise da regressão pode ser usada como um método descritivo da análise de dados (como, por exemplo, o ajustamento de curvas) sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados. Regressão designa também uma equação matemática que descreva a relação entre duas ou mais variáveis

Quando há apenas 2 variáveis em análise diz-se "regressão simples" e quando há mais do que 2 variáveis envolvidas, diz-se "regressão múltipla". Na fórmula original tem-se

y = Xb + e

, que deve ser lido como o parametro matricial b é o que ajusta os y e os x observados.

A fórmula de cálculo de b é b = (X'X)^-1 X'y, onde X' é a matriz transposta de X e X^-1 é a matriz inversa de X.

Acompanhe o exemplo

Suponhamos querer estudar o comportamento do consumo de bens em uma determinada família, sujeita a dois parâmetros:

- renda e
- taxa de juros

Dos estudos de econometria, sabe-se que o consumo aumenta em função da renda e diminui em função da taxa de juros.

O trabalho começa pelo levantamento de dados prévios e obteve-se a seguinte tabela:

medida	consumo y	renda x1	tx juros x2
1	122	139	11,5%
2	114	126	12,0%
3	86	90	10,5%
4	134	144	9,0%
5	146	163	10,0%
6	107	136	12,0%
7	68	61	10,5%
8	117	62	8,0%
9	71	41	10,0%
10	98	120	11,5%

A matriz X do nosso problema é:

1	139	0.115
1	126	0.12
1	90	0.105
1	144	0.09
1	163	0.1
1	136	0.12
1	61	0.105
1	62	0.08
1	41	0.1
1	120	0.115

Note que a primeira coluna sempre é formada de unidades e as demais colunas são as variáveis em estudo (no caso: renda e taxa de juros).

Tomando a matriz transposta fica-se com

1	1	1	1	1	1	1	1	1	1
139	126	90	144	163	136	61	62	41	120
0.11	0.12	0.10	0.09	0.1	0.12	0.10	0.08	0.1	0.11

Multiplicando uma pela outra, tem-se

10	1082	1.05
1082	132744	115.4
1.05	115.4	0.1118

E a matriz inversa desta é

7.225436859	0.00095983853	-68.85039418
0.00095983853	0.00007350542	-0.084886905
-68.85039418	-0.08488690598	743.1919753

Agora, multiplicando a inversa por X' fica:

-5	-.91	.08	1.16	.49	-.90	.05	1.7	.3	-.5
.0	.00	-.00	.00	.00	.0	-.00	-.0	-.0	.0
4.8	9.63	1.54	-14.1	-8.36	8.7	4.0	-14.6	1.9	6.4

Agora multiplicando esta matriz pela matriz coluna de y (consumo) que é (devidamente horizontalizada para economizar papel)

122	114	86	134	146	107	68	117	71	98
-----	-----	----	-----	-----	-----	----	-----	----	----

Fica

148.5220194
0.6135933813
-1034.40784

Interpretação

Neste problema, o consumo é

Consumo =

148,52 + 0.6135 x Renda - 1.034.41 x TxJuros

Ou seja, tirando a parte do Consumo que não é influenciada pela Taxa de Juros, o incremento de \$ 1 na Renda causa um incremento esperado de \$ 0,6136 no Consumo; além disso, o incremento de 1 ponto percentual (0,01) na Taxa de Juros causa um decréscimo esperado de \$ 10,3441 no Consumo.

Para você fazer

1. Considere o mesmo problema com os seguintes valores obtidos

medida	consumo y	renda x1	tx juros x2
1	98	123	.1030
2	91	115	.1050
3	105	108	.1080
4	95	112	.1030
5	90	128	.1060
6	88	108	.1080
7	109	124	.1110
8	104	128	.1050
9	107	118	.1140
10	82	102	.1160

Ache os parâmetros de renda e taxa de juros para o problema proposto.

t. fixo	renda	tx de juros
---------	-------	-------------

2. Mais um problema, agora envolvendo um outro modelo econométrico (obviamente FICTICIO) no qual estão relacionadas 3 variáveis independentes e uma dependente. Não se esqueça que a matemática lava as mãos sobre essas definições. É coisa do planejador. São elas:

t	dep.	consumo de whisky importado (em caixas)
x	indep.	venda de propriedades em Caiobá (em um)
y	indep.	número de falências (sentido oposto, em um)
z	indep.	número de BMW licenciadas (em um)

O objetivo então é achar

t = a + b.x + c.y + d.z

No último ano foram levantados os valores de x, y e z e obteve-se

medida	t	x	y	z
1	-2102899	473	746	-33
2	4603229	807	237	848
3	1782902	468	370	512
4	967962	625	516	374
5	-483649	187	904	667
6	516299	247	426	397
7	670004	44	541	663
8	637353	833	377	40
9	1314894	877	419	215
10	394546	384	401	265

Usando a teoria acima, você deve achar os valores de a, b, c e d.

a	b	c	d
---	---	---	---

3. O problema agora é que você é o dono do maior restaurante de Santa Felicidade. Precisa prever a quantidade de refeições a serem servidas no próximo Dia das Mães. Depois de muito estudar, você contratou um engenheiro que levantou dados dos últimos 6 anos e elaborou o seguinte modelo econométrico:

a	indep	número de refeições a serem servidas no próximo dia das mães
b	dep	número de casamentos em Curitiba no ano anterior
c	dep	renda média das famílias no Paraná no ano anterior, em reais
d	dep	frangos abatidos na última semana do mês de abril no Paraná
e	dep	veículos emplacadas na ciretran de S. Felicidade no ano anterior

A fórmula do modelo é:

a = α + βb + γc + δd + ε

Os dados:

	a	b	c	d	e
1	7806	11515	7201	119138	5272
2	7411	12122	8246	113238	3360
3	8957	14446	9599	138124	3521
4	11180	16915	9063	168518	10761
5	10425	16996	7298	156647	10657
6	6914	19628	8545	99438	4682

Achar os valores de α, β, γ, δ e ε.

α	β	γ	δ	ε
---	---	---	---	---

Observação importante: No Python, ao gerar as matrizes, não esquecer de pedir campos float, porque senão ele pode resolver usando inteiros e vai achar respostas erradas...



122-69039 - f/na

Regressão múltipla

Quando o passado permite prever o futuro, há um método estatístico denominado Regressão Linear que busca o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Em estatística, regressão é uma técnica que permite explorar e inferir a relação de uma variável dependente (variável de resposta) com variáveis independentes específicas (variáveis explicatórias). A análise da regressão pode ser usada como um método descritivo da análise de dados (como, por exemplo, o ajustamento de curvas) sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados. Regressão designa também uma equação matemática que descreva a relação entre duas ou mais variáveis

Quando há apenas 2 variáveis em análise diz-se "regressão simples"e quando há mais do que 2 variáveis envolvidas, diz-se "regressão múltipla".

Na fórmula original tem-se

y = Xb + ε

, que deve ser lido como o parametro matricial b é o que ajusta os y e os x observados.

A fórmula de cálculo de b é b = (X'X)⁻¹X'y, onde X' é a matriz transposta de X é X⁻¹ é a matriz inversa de X.

Acompanhe o exemplo

Suponhamos querer estudar o comportamento do consumo de bens em uma determinada família, sujeita a dois parâmetros:

- renda e
- taxa de juros

Dos estudos de econometria, sabe-se que o consumo aumenta em função da renda e diminui em função da taxa de juros.

O trabalho começa pelo levantamento de dados prévios e obteve-se a seguinte tabela:

medida	consumo y	renda x ₁	tx juros x ₂
1	122	139	11,5%
2	114	126	12,0%
3	86	90	10,5%
4	134	144	9,0%
5	146	163	10,0%
6	107	136	12,0%
7	68	61	10,5%
8	117	62	8,0%
9	71	41	10,0%
10	98	120	11,5%

A matriz X do nosso problema é:

1	139	0.115
1	126	0.12
1	90	0.105
1	144	0.09
1	163	0.1
1	136	0.12
1	61	0.105
1	62	0.08
1	41	0.1
1	120	0.115

Note que a primeira coluna sempre é formada de unidades e as demais colunas são as variáveis em estudo (no caso: renda e taxa de juros).

Tomando a matriz transposta fica-se com

1	1	1	1	1	1	1	1	1	1
139	126	90	144	163	136	61	62	41	120
0.11	0.12	0.10	0.09	0.1	0.12	0.10	0.08	0.1	0.11

Multiplicando uma pela outra, tem-se

10	1082	1.05
1082	132744	115.4
1.05	115.4	0.1118

E a matriz inversa desta é

7.225436859	0.00095983853	-68.85039418
0.00095983853	0.0007350542	-0.084886905
-68.85039418	-0.08488690598	743.1919753

Agora, multiplicando a inversa por X' fica:

-5	-.91	.08	1.16	.49	-.90	.05	1.7	.3	-.5
.0	.00	-.00	.00	.00	.0	-.00	-.0	-.0	.0
4.8	9.63	1.54	-14.1	-8.36	8.7	4.0	-14.6	1.9	6.4

Agora multiplicando esta matriz pela matriz coluna de y (consumo) que é (devidamente horizontalizada para economizar papel)

122	114	86	134	146	107	68	117	71	98
-----	-----	----	-----	-----	-----	----	-----	----	----

Fica

148.5220194
0.6135933813
-1034.40784

Interpretação

Neste problema, o consumo é

Consumo =

148,52 + 0.6135 × Renda – 1.034.41 × TxJuros

Ou seja, tirando a parte do Consumo que não é influenciada pela Taxa de Juros, o incremento de \$ 1 na Renda causa um incremento esperado de \$ 0,6136 no Consumo; além disso, o incremento de 1 ponto percentual (0,01) na Taxa de Juros causa um decréscimo esperado de \$ 10,3441 no Consumo.

Para você fazer

1. Considere o mesmo problema com os seguintes valores obtidos

medida	consumo y	renda x ₁	tx juros x ₂
1	92	111	.1190
2	92	124	.1020
3	105	120	.0950
4	92	128	.1090
5	101	103	.1040
6	98	103	.1030
7	83	104	.0970
8	108	123	.1140
9	82	123	.1100
10	96	101	.1170

Ache os parâmetros de renda e taxa de juros para o problema proposto.

t. fixo	renda	tx de juros
---------	-------	-------------

2. Mais um problema, agora envolvendo um outro modelo econométrico (obviamente FICTICIO) no qual estão relacionadas 3 variáveis independentes e uma dependente. Não se esqueça que a matemática lava as mãos sobre essas definições. É coisa do planejador. São elas:

t	dep.	consumo de whisky importado (em caixas)
x	indep.	venda de propriedades em Caiobá (em um)
y	indep.	número de falências (sentido oposto, em um)
z	indep.	número de BMW licenciadas (em um)

O objetivo então é achar

t = a + b.x + c.y + d.z

No último ano foram levantados os valores de x, y e z e obteve-se

medida	t	x	y	z
1	896182	10	193	388
2	420057	265	884	668
3	1987823	741	551	331
4	1400615	-85	382	770
5	2351763	13	128	689
6	392393	545	334	-88
7	-384467	237	429	80
8	2893364	656	493	580
9	3191512	648	723	870
10	-197476	254	358	47

Usando a teoria acima, você deve achar os valores de a, b, c e d.

a	b	c	d
---	---	---	---

3. O problema agora é que você é o dono do maior restaurante de Santa Felicidade. Precisa prever a quantidade de refeições a serem servidas no próximo Dia das Mães. Depois de muito estudar, você contratou um engenheiro que levantou dados dos últimos 6 anos e elaborou o seguinte modelo econométrico:

a	indep	número de refeições a serem servidas no próximo dia das mães
b	dep	número de casamentos em Curitiba no ano anterior
c	dep	renda média das famílias no Paraná no ano anterior, em reais
d	dep	frangos abatidos na última semana do mês de abril no Paraná
e	dep	veículos emplacadas na ciretran de S. Felicidade no ano anterior

A fórmula do modelo é:

a = α + βb + γc + δd + ε

Os dados:

	a	b	c	d	e
1	9324	14135	13891	106325	10449
2	9399	14425	11150	110628	9563
3	10349	14919	13424	125390	7140
4	12384	15340	10272	158741	7518
5	9869	19577	10425	115741	8000
6	9841	19709	6841	119915	6597

Achar os valores de α, β, γ, δ e ε.

α	β	γ	δ	ε
---	---	---	---	---

Observação importante: No Python, ao gerar as matrizes, não esquecer de pedir campos float, porque senão ele pode resolver usando inteiros e vai achar respostas erradas...



122-69046 - f/na

Regressão múltipla

Quando o passado permite prever o futuro, há um método estatístico denominado Regressão Linear que busca o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Em estatística, regressão é uma técnica que permite explorar e inferir a relação de uma variável dependente (variável de resposta) com variáveis independentes específicas (variáveis explicatórias). A análise da regressão pode ser usada como um método descritivo da análise de dados (como, por exemplo, o ajustamento de curvas) sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados. Regressão designa também uma equação matemática que descreva a relação entre duas ou mais variáveis

Quando há apenas 2 variáveis em análise diz-se "regressão simples" e quando há mais do que 2 variáveis envolvidas, diz-se "regressão múltipla". Na fórmula original tem-se

y = Xb + e

, que deve ser lido como o parametro matricial b é o que ajusta os y e os x observados.

A fórmula de cálculo de b é b = (X'X)^-1 X'y, onde X' é a matriz transposta de X e X^-1 é a matriz inversa de X.

Acompanhe o exemplo

Suponhamos querer estudar o comportamento do consumo de bens em uma determinada família, sujeita a dois parâmetros:

- renda e
- taxa de juros

Dos estudos de econometria, sabe-se que o consumo aumenta em função da renda e diminui em função da taxa de juros.

O trabalho começa pelo levantamento de dados prévios e obteve-se a seguinte tabela:

medida	consumo y	renda x1	tx juros x2
1	122	139	11,5%
2	114	126	12,0%
3	86	90	10,5%
4	134	144	9,0%
5	146	163	10,0%
6	107	136	12,0%
7	68	61	10,5%
8	117	62	8,0%
9	71	41	10,0%
10	98	120	11,5%

A matriz X do nosso problema é:

1	139	0.115
1	126	0.12
1	90	0.105
1	144	0.09
1	163	0.1
1	136	0.12
1	61	0.105
1	62	0.08
1	41	0.1
1	120	0.115

Note que a primeira coluna sempre é formada de unidades e as demais colunas são as variáveis em estudo (no caso: renda e taxa de juros).

Tomando a matriz transposta fica-se com

1	1	1	1	1	1	1	1	1	1
139	126	90	144	163	136	61	62	41	120
0.11	0.12	0.10	0.09	0.1	0.12	0.10	0.08	0.1	0.11

Multiplicando uma pela outra, tem-se

10	1082	1.05
1082	132744	115.4
1.05	115.4	0.1118

E a matriz inversa desta é

7.225436859	0.00095983853	-68.85039418
0.00095983853	0.00007350542	-0.084886905
-68.85039418	-0.08488690598	743.1919753

Agora, multiplicando a inversa por X' fica:

-5	-.91	.08	1.16	.49	-.90	.05	1.7	.3	-.5
.0	.00	-.00	.00	.00	.0	-.00	-.0	-.0	.0
4.8	9.63	1.54	-14.1	-8.36	8.7	4.0	-14.6	1.9	6.4

Agora multiplicando esta matriz pela matriz coluna de y (consumo) que é (devidamente horizontalizada para economizar papel)

122	114	86	134	146	107	68	117	71	98
-----	-----	----	-----	-----	-----	----	-----	----	----

Fica

148.5220194
0.6135933813
-1034.40784

Interpretação

Neste problema, o consumo é

Consumo =

148,52 + 0.6135 x Renda - 1.034.41 x TxJuros

Ou seja, tirando a parte do Consumo que não é influenciada pela Taxa de Juros, o incremento de \$ 1 na Renda causa um incremento esperado de \$ 0,6136 no Consumo; além disso, o incremento de 1 ponto percentual (0,01) na Taxa de Juros causa um decréscimo esperado de \$ 10,3441 no Consumo.

Para você fazer

1. Considere o mesmo problema com os seguintes valores obtidos

medida	consumo y	renda x1	tx juros x2
1	106	124	.1060
2	106	106	.0940
3	92	125	.1010
4	97	109	.0910
5	102	122	.1060
6	88	117	.1040
7	96	127	.1110
8	88	128	.0990
9	88	126	.1080
10	109	109	.0980

Ache os parâmetros de renda e taxa de juros para o problema proposto.

t. fixo	renda	tx de juros
---------	-------	-------------

2. Mais um problema, agora envolvendo um outro modelo econométrico (obviamente FICTICIO) no qual estão relacionadas 3 variáveis independentes e uma dependente. Não se esqueça que a matemática lava as mãos sobre essas definições. É coisa do planejador. São elas:

t	dep.	consumo de whisky importado (em caixas)
x	indep.	venda de propriedades em Caiobá (em um)
y	indep.	número de falências (sentido oposto, em um)
z	indep.	número de BMW licenciadas (em um)

O objetivo então é achar

t = a + b.x + c.y + d.z

No último ano foram levantados os valores de x, y e z e obteve-se

medida	t	x	y	z
1	-839009	491	362	-58
2	352871	273	630	536
3	2546054	890	345	820
4	314890	165	859	725
5	3260047	878	11	767
6	12090	395	770	523
7	-1123415	-97	591	156
8	-591766	429	526	153
9	1845202	528	400	734
10	-1444999	620	907	167

Usando a teoria acima, você deve achar os valores de a, b, c e d.

a	b	c	d
---	---	---	---

3. O problema agora é que você é o dono do maior restaurante de Santa Felicidade. Precisa prever a quantidade de refeições a serem servidas no próximo Dia das Mães. Depois de muito estudar, você contratou um engenheiro que levantou dados dos últimos 6 anos e elaborou o seguinte modelo econométrico:

a	indep	número de refeições a serem servidas no próximo dia das mães
b	dep	número de casamentos em Curitiba no ano anterior
c	dep	renda média das famílias no Paraná no ano anterior, em reais
d	dep	frangos abatidos na última semana do mês de abril no Paraná
e	dep	veículos emplacadas na ciretran de S. Felicidade no ano anterior

A fórmula do modelo é:

a = α + βb + γc + δd + ε

Os dados:

	a	b	c	d	e
1	9730	12346	12714	139192	6390
2	9381	12737	8748	137844	4256
3	10873	13426	6997	158603	7603
4	10409	14986	7504	146394	10585
5	8648	15513	8651	124299	4627
6	10373	18466	9242	149463	5969

Achar os valores de α, β, γ, δ e ε.

α	β	γ	δ	ε
---	---	---	---	---

Observação importante: No Python, ao gerar as matrizes, não esquecer de pedir campos float, porque senão ele pode resolver usando inteiros e vai achar respostas erradas...



122-69060 - f/na

Regressão múltipla

Quando o passado permite prever o futuro, há um método estatístico denominado Regressão Linear que busca o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Em estatística, regressão é uma técnica que permite explorar e inferir a relação de uma variável dependente (variável de resposta) com variáveis independentes específicas (variáveis explicatórias). A análise da regressão pode ser usada como um método descritivo da análise de dados (como, por exemplo, o ajustamento de curvas) sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados. Regressão designa também uma equação matemática que descreva a relação entre duas ou mais variáveis

Quando há apenas 2 variáveis em análise diz-se "regressão simples" e quando há mais do que 2 variáveis envolvidas, diz-se "regressão múltipla".

Na fórmula original tem-se

y = Xb + ε

, que deve ser lido como o parametro matricial b é o que ajusta os y e os x observados.

A fórmula de cálculo de b é b = (X'X)⁻¹X'y, onde X' é a matriz transposta de X é X⁻¹ é a matriz inversa de X.

Acompanhe o exemplo

Suponhamos querer estudar o comportamento do consumo de bens em uma determinada família, sujeita a dois parâmetros:

- renda e
- taxa de juros

Dos estudos de econometria, sabe-se que o consumo aumenta em função da renda e diminui em função da taxa de juros.

O trabalho começa pelo levantamento de dados prévios e obteve-se a seguinte tabela:

medida	consumo y	renda x ₁	tx juros x ₂
1	122	139	11,5%
2	114	126	12,0%
3	86	90	10,5%
4	134	144	9,0%
5	146	163	10,0%
6	107	136	12,0%
7	68	61	10,5%
8	117	62	8,0%
9	71	41	10,0%
10	98	120	11,5%

A matriz X do nosso problema é:

1	139	0.115
1	126	0.12
1	90	0.105
1	144	0.09
1	163	0.1
1	136	0.12
1	61	0.105
1	62	0.08
1	41	0.1
1	120	0.115

Note que a primeira coluna sempre é formada de unidades e as demais colunas são as variáveis em estudo (no caso: renda e taxa de juros).

Tomando a matriz transposta fica-se com

1	1	1	1	1	1	1	1	1	1
139	126	90	144	163	136	61	62	41	120
0.11	0.12	0.10	0.09	0.1	0.12	0.10	0.08	0.1	0.11

Multiplicando uma pela outra, tem-se

10	1082	1.05
1082	132744	115.4
1.05	115.4	0.1118

E a matriz inversa desta é

7.225436859	0.00095983853	-68.85039418
0.00095983853	0.00007350542	-0.084886905
-68.85039418	-0.08488690598	743.1919753

Agora, multiplicando a inversa por X' fica:

-5	-.91	.08	1.16	.49	-.90	.05	1.7	.3	-.5
.0	.00	-.00	.00	.00	.0	-.00	-.0	-.0	.0
4.8	9.63	1.54	-14.1	-8.36	8.7	4.0	-14.6	1.9	6.4

Agora multiplicando esta matriz pela matriz coluna de y (consumo) que é (devidamente horizontalizada para economizar papel)

122	114	86	134	146	107	68	117	71	98
-----	-----	----	-----	-----	-----	----	-----	----	----

Fica

148.5220194
0.6135933813
-1034.40784

Interpretação

Neste problema, o consumo é

Consumo =

148,52 + 0.6135 × Renda - 1.034.41 × TxJuros

Ou seja, tirando a parte do Consumo que não é influenciada pela Taxa de Juros, o incremento de \$ 1 na Renda causa um incremento esperado de \$ 0,6136 no Consumo; além disso, o incremento de 1 ponto percentual (0,01) na Taxa de Juros causa um decréscimo esperado de \$ 10,3441 no Consumo.

Para você fazer

1. Considere o mesmo problema com os seguintes valores obtidos

medida	consumo y	renda x ₁	tx juros x ₂
1	96	120	.1150
2	109	106	.0910
3	83	120	.1190
4	99	108	.0920
5	108	114	.1160
6	98	112	.1190
7	89	105	.1200
8	81	128	.0930
9	102	118	.0970
10	87	116	.1040

Ache os parâmetros de renda e taxa de juros para o problema proposto.

t. fixo	renda	tx de juros
---------	-------	-------------

2. Mais um problema, agora envolvendo um outro modelo econométrico (obviamente FICTICIO) no qual estão relacionadas 3 variáveis independentes e uma dependente. Não se esqueça que a matemática lava as mãos sobre essas definições. É coisa do planejador. São elas:

t	dep.	consumo de whisky importado (em caixas)
x	indep.	venda de propriedades em Caiobá (em um)
y	indep.	número de falências (sentido oposto, em um)
z	indep.	número de BMW licenciadas (em um)

O objetivo então é achar

t = a + b.x + c.y + d.z

No último ano foram levantados os valores de x, y e z e obteve-se

medida	t	x	y	z
1	1416500	742	400	205
2	-3481238	-36	728	125
3	815670	390	915	790
4	3364449	637	439	629
5	-2564900	-12	419	-28
6	1199577	651	382	202
7	-615312	512	413	3
8	5591886	668	174	735
9	-795974	612	928	407
10	2352663	305	406	612

Usando a teoria acima, você deve achar os valores de a, b, c e d.

a	b	c	d
---	---	---	---

3. O problema agora é que você é o dono do maior restaurante de Santa Felicidade. Precisa prever a quantidade de refeições a serem servidas no próximo Dia das Mães. Depois de muito estudar, você contratou um engenheiro que levantou dados dos últimos 6 anos e elaborou o seguinte modelo econométrico:

a	indep	número de refeições a serem servidas no próximo dia das mães
b	dep	número de casamentos em Curitiba no ano anterior
c	dep	renda média das famílias no Paraná no ano anterior, em reais
d	dep	frangos abatidos na última semana do mês de abril no Paraná
e	dep	veículos emplacadas na ciretran de S. Felicidade no ano anterior

A fórmula do modelo é:

a = α + βb + γc + δd + ε

Os dados:

	a	b	c	d	e
1	10492	11181	12560	166003	9011
2	7486	11604	7535	110636	10373
3	6992	13614	13514	99290	7133
4	6162	17747	8752	90391	4060
5	8972	18164	10319	132419	10305
6	8169	19756	7836	129583	3464

Achar os valores de α, β, γ, δ e ε.

α	β	γ	δ	ε
---	---	---	---	---

Observação importante: No Python, ao gerar as matrizes, não esquecer de pedir campos float, porque senão ele pode resolver usando inteiros e vai achar respostas erradas...



122-69077 - f/na

Regressão múltipla

Quando o passado permite prever o futuro, há um método estatístico denominado Regressão Linear que busca o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Em estatística, regressão é uma técnica que permite explorar e inferir a relação de uma variável dependente (variável de resposta) com variáveis independentes específicas (variáveis explicatórias). A análise da regressão pode ser usada como um método descritivo da análise de dados (como, por exemplo, o ajustamento de curvas) sem serem necessárias quaisquer suposições acerca dos processos que permitiram gerar os dados. Regressão designa também uma equação matemática que descreva a relação entre duas ou mais variáveis

Quando há apenas 2 variáveis em análise diz-se "regressão simples" e quando há mais do que 2 variáveis envolvidas, diz-se "regressão múltipla". Na fórmula original tem-se

y = Xb + e

, que deve ser lido como o parametro matricial b é o que ajusta os y e os x observados.

A fórmula de cálculo de b é b = (X'X)^-1 X'y, onde X' é a matriz transposta de X e X^-1 é a matriz inversa de X.

Acompanhe o exemplo

Suponhamos querer estudar o comportamento do consumo de bens em uma determinada família, sujeita a dois parâmetros:

- renda e
- taxa de juros

Dos estudos de econometria, sabe-se que o consumo aumenta em função da renda e diminui em função da taxa de juros.

O trabalho começa pelo levantamento de dados prévios e obteve-se a seguinte tabela:

medida	consumo y	renda x1	tx juros x2
1	122	139	11,5%
2	114	126	12,0%
3	86	90	10,5%
4	134	144	9,0%
5	146	163	10,0%
6	107	136	12,0%
7	68	61	10,5%
8	117	62	8,0%
9	71	41	10,0%
10	98	120	11,5%

A matriz X do nosso problema é:

1	139	0.115
1	126	0.12
1	90	0.105
1	144	0.09
1	163	0.1
1	136	0.12
1	61	0.105
1	62	0.08
1	41	0.1
1	120	0.115

Note que a primeira coluna sempre é formada de unidades e as demais colunas são as variáveis em estudo (no caso: renda e taxa de juros).

Tomando a matriz transposta fica-se com

1	1	1	1	1	1	1	1	1	1
139	126	90	144	163	136	61	62	41	120
0.11	0.12	0.10	0.09	0.1	0.12	0.10	0.08	0.1	0.11

Multiplicando uma pela outra, tem-se

10	1082	1.05
1082	132744	115.4
1.05	115.4	0.1118

E a matriz inversa desta é

7.225436859	0.00095983853	-68.85039418
0.00095983853	0.00007350542	-0.084886905
-68.85039418	-0.08488690598	743.1919753

Agora, multiplicando a inversa por X' fica:

-5	-.91	.08	1.16	.49	-.90	.05	1.7	.3	-.5
.0	.00	-.00	.00	.00	.0	-.00	-.0	-.0	.0
4.8	9.63	1.54	-14.1	-8.36	8.7	4.0	-14.6	1.9	6.4

Agora multiplicando esta matriz pela matriz coluna de y (consumo) que é (devidamente horizontalizada para economizar papel)

122	114	86	134	146	107	68	117	71	98
-----	-----	----	-----	-----	-----	----	-----	----	----

Fica

148.5220194	0.6135933813
-1034.40784	

Interpretação

Neste problema, o consumo é

Consumo =

148,52 + 0.6135 x Renda - 1.034.41 x TxJuros

Ou seja, tirando a parte do Consumo que não é influenciada pela Taxa de Juros, o incremento de \$ 1 na Renda causa um incremento esperado de \$ 0,6136 no Consumo; além disso, o incremento de 1 ponto percentual (0,01) na Taxa de Juros causa um decréscimo esperado de \$ 10,3441 no Consumo.

Para você fazer

1. Considere o mesmo problema com os seguintes valores obtidos

medida	consumo y	renda x1	tx juros x2
1	89	125	.1080
2	81	115	.1020
3	110	116	.0920
4	102	111	.1110
5	102	101	.0990
6	89	129	.1190
7	97	113	.1090
8	101	105	.1060
9	104	122	.1090
10	100	106	.1120

Ache os parâmetros de renda e taxa de juros para o problema proposto.

t. fixo	renda	tx de juros
---------	-------	-------------

2. Mais um problema, agora envolvendo um outro modelo econométrico (obviamente FICTICIO) no qual estão relacionadas 3 variáveis independentes e uma dependente. Não se esqueça que a matemática lava as mãos sobre essas definições. É coisa do planejador. São elas:

t	dep.	consumo de whisky importado (em caixas)
x	indep.	venda de propriedades em Caiobá (em um)
y	indep.	número de falências (sentido oposto, em um)
z	indep.	número de BMW licenciadas (em um)

O objetivo então é achar

t = a + b.x + c.y + d.z

No último ano foram levantados os valores de x, y e z e obteve-se

medida	t	x	y	z
1	-909941	386	594	298
2	-722689	682	356	45
3	352229	277	381	388
4	-1158635	479	840	464
5	678035	814	55	29
6	-2953695	135	780	87
7	-668295	807	803	466
8	-1632703	829	901	347
9	1670305	113	79	415
10	-2800447	667	703	-66

Usando a teoria acima, você deve achar os valores de a, b, c e d.

a	b	c	d
---	---	---	---

3. O problema agora é que você é o dono do maior restaurante de Santa Felicidade. Precisa prever a quantidade de refeições a serem servidas no próximo Dia das Mães. Depois de muito estudar, você contratou um engenheiro que levantou dados dos últimos 6 anos e elaborou o seguinte modelo econométrico:

a	indep	número de refeições a serem servidas no próximo dia das mães
b	dep	número de casamentos em Curitiba no ano anterior
c	dep	renda média das famílias no Paraná no ano anterior, em reais
d	dep	frangos abatidos na última semana do mês de abril no Paraná
e	dep	veículos emplacadas na ciretran de S. Felicidade no ano anterior

A fórmula do modelo é:

a = α + βb + γc + δd + ε

Os dados:

	a	b	c	d	e
1	5342	11586	6804	124477	6731
2	6423	13264	7970	150743	7839
3	5381	13709	10347	117711	9845
4	6021	15547	9382	141632	6219
5	5251	16230	13490	119120	5969
6	4255	16386	12380	92239	6241

Achar os valores de α, β, γ, δ e ε.

α	β	γ	δ	ε
---	---	---	---	---

Observação importante: No Python, ao gerar as matrizes, não esquecer de pedir campos float, porque senão ele pode resolver usando inteiros e vai achar respostas erradas...



122-69796 - f/na