



Prof. Landeira-Fernandez

Bioestatística

**Esta aula foi compilada por alunos. Caso encontre erros,
favor procurar no email: luisfca@gmail.com ou
landeira@puc-rio.br**

Rio de Janeiro, 23 de junho de 2015.

AULA 9: DELINEAMENTO CORRELACIONALS

Um tipo de pesquisa muito utilizado em ciências humanas e sociais é aquela em que vc possui um único grupo e muitas medidas.

Neste caso ainda não esta claro as possíveis relações de causa e efeito, por essa razão existem muitas medidas, neste caso utiliza-se técnicas de correlação e regressão partindo-se do principio que a relação entre as variáveis é de natureza linear.

Coeficiente de correlação

Este coeficiente sempre tenta descobrir a força de associação e a natureza dessa associação entre 2 variaveis. A correlação (r) pode variar de 0 a 1. Quanto mais próximo de 1 mais forte é essa correlação.

0,03 – fraca, 0,3 a 0,6 média, 0,6 a 0,9 forte

O coeficiente de correlação quando elevado ao quadrado serve como indicador do quanto uma variável consegue explicar o comportamento da outra variável, por exemplo, uma correlação de 0,3 produz R^2 de 0,09 ou seja uma variável EXPLICA (é responsável) po r10% do comportamento da outra variável.

Por outro lado se uma correlação é igual a 0,7, gerando R^2 de 0,49 o impacto de uma variável sobre o comportamento da outra é de 50%.

A NATUREZA DA CORRELAÇÃO

Alem do valor numérico, o coeficiente de correlação indica também que tipo de associação existe entre 2 variáveis. Isso é dado através do sinal. Se o sinal for positivo, a correlação é diretamente proporcional, se o sinal for negativo é inversamente proporcional.

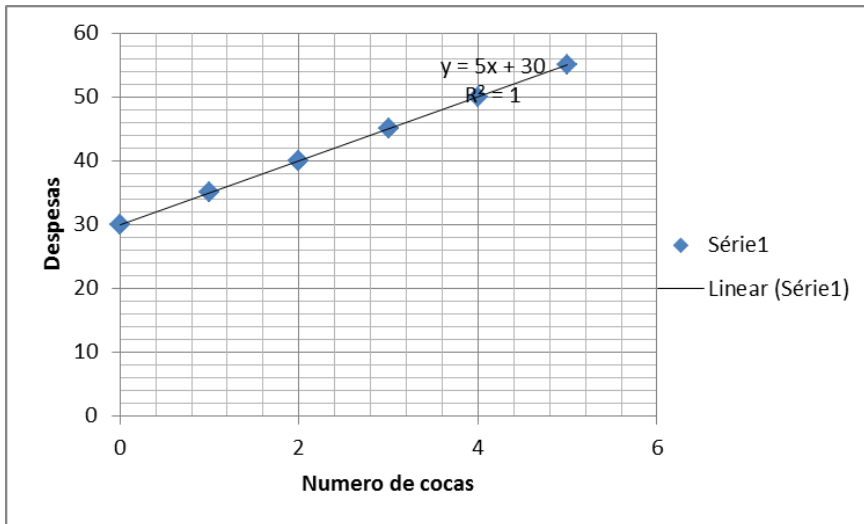
REGRESSÃO

Um pequeno exercício pode ilustrar muito bem o conceito de regressão.



Vc vai a um bar assistir um show de jazz. O couvert artístico é de R\$30,00. Voce so bebe coca cola. O preço da coca é R\$5,00. O Bar não cobra 10%. Qual o valor da sua conta? R\$50,00

Indique a $V_i(x)$ e $V_d(y)$ dessa situação: $Y = aX+b$



a= preço da coca cola

b= couvert

x => variável (numero de cocas)

Na regressão linear a V_d , que é aquilo que se quer explicar, recebe a designação de “Y”. A V_i por sua vez, que é uma possível variável que supostamente causa V_d e que se quer testar, recebe a denominação de “x”.

A relação entre essas duas variáveis se dá através da função da reta que busca explicar a correlação entre as duas variáveis. A função é $y'xa+b$, onde y é a V_d , x é a V_i , a é a inclinação da reta, e b é o intercepto.

No nosso exercício, a conta y =numero de cocas consumido (x) multiplicado pelo valor da coca (a que é constante) mais o valor do couvert (b que é constante).

REGRESSAO MULTIPLA

A regressão linear simples incorpora apenas uma única V_i . Podemos testar situações onde existem duas ou mais V_i 's, ou seja o comportamento da V_d Y pode ser determinado por mais de uma V_i (x_1, x_2, x_3, x_n).

Continuando o exercício do bar.



Voce vai assistir o show com um colega, que além de coca também toma cerveja (R\$7,00). Qual o valor da conta do seu amigo?

Conforme pode se observar no exemplo o que determinou a conta do colega, foi algo mais complexo, além da coca, consumiu cerveja, portanto existe 2 Vi's. Pode-se agora expressar a regressão múltipla, associando cada Vi a sua respectiva constante da seguinte forma

$$Y = x_1a_1 + x_2a_2 + b$$

Análise dos dados

A regressão de uma Vd pode ser testada no SPSS definindo-se uma única Vd e várias Vi's. O valor de cada constante associada a regressão múltipla se dá através de Beta.

Cada beta pode ser testado na sua significância através de um teste T.

Nosso caso que foi preparado artificialmente sabemos que a conta e função do consumo de cerveja e coca mas não da temperatura ou da musica (certamente a temperatura e a musica podem influenciar o consumo de cerveja e de coca e indiretamente afeta o valor da conta mas certamente a conta não varia através de uma relação CAUSAL da temperatura ambiente ou da musica.)

Essas hipóteses podem ser testadas através de uma análise de regressão. Sabendo-se os valores das constantes pode-se então prever o comportamento futuro da Vd.

Criticas ao método correlacional

Existe uma correlação alta entre ansiedade e depressão, quem causa quem?

Uma alta correlação não indica como se dá essa relação de causa e efeito, no entanto não é possível dizer que é Vi e Vd. Entretanto existe uma critica mais severa, que é o da terceira variável, que não é do conhecimento do pesquisador mas que mantém relação de causa e efeito com as 2 variáveis que estão sendo estudadas.

Existe uma correlação alta entre consumo de sorvete e assalto a mão armada. $R=0,6$

Isso nos leva a uma das principais limitações bem reconhecidas: "Uma alta correlação não indica necessariamente em uma relação de causa e efeito.

Deve-se notar que mesmo que não exista a relação de causa e efeito entre as duas variáveis, é possível fazer a regressão estabelecendo as respectivas constantes da função da reta e assim prever o comportamento futuro da Vd, desde que a correlação exista e os betas sejam significativos.

Topicos abordados na disciplina



Aspectos filosoficos

Se alguma coisa existe ela pode ser medida, a medida é sempre é indireta através de instrumentos.

Psicometria – è possivel medir um aspecto psicologico? As medidas não são do tipo tudo ou nada, existem gradações. A forma de avaliar a qualidade de uma medida é através da sua validade.

Medidas em psicologia não tem validade. Não tem fundamentação teorica robusta para validar o fenomeno.

Técnicas estatísticas para avalias a qualidade psicometrica de uma medida

- 1- Fidedignidade (alfa de cronbach)
- 2- Validade (analise fatorial) construto teorico

Um aspecto importante na analise de um instrumento é analise de itens (correlação item total)

O emprego clinico de um instrumento (avaliar a capacidade do instrumento de gerar diagnostico)

Especificidade e do diagnostico

Estudo de relações de causa e efeito – Uso de instrumentos para testar teorias. Relações de causa e efeito entre variaveis (V_i e V_d)

Estudo experimental (2 grupos – grupo teste e grupo controle)

Comparação de 2 grupos o que determina o numero de grupo s são os níveis de V_1 (ANOVA 1 via) ou numero de V_i ' (ANOVA de x vias)

Os fenomenos psicologicos não são uni causais, pois existem interação entre diferentes V_i 's portanto ANOVA de x vias sempre procura INTERAÇÃO.

Experimental (Melhor delineamento) De longe o melhor para testar relações de causa e efeito, questões éticas impedem manipular V_i

Quase experimental – Selecciona ao inves de manipular a V_i . Limitações, ao seleccionar V_i pode causar viéz.

Para tratar viez que são detectados na pesquisa podem ser controlados através da ANCOVA

2 delinamentos : **Inter** e intra sujeito

Inter evita contaminação mas o efeito tem que ser de alta magnitude



Intra (sujeito com seu próprio controle) o poder da análise é maior, a limitação o sujeito está sujeito a contaminação.

Metodo experimental e quase experimental testa hipóteses muito bem formuladas dai o numero reduzidos de Vd's, sendo assim um número grande de níveis de Vi's ou diversas Vi's.

Metodo correlacional não sabe o que causa o que mas tem muitas medidas

Pesquisa correlacional. Apresenta um único grupo com várias medidas onde se busca correlaciona pares de variáveis. Quanto maior a correlação mais forte a relação entre as variáveis.

A regressão permite testar a significância de uma Vi sobre a Vd. Mesmo sendo significativa não indica necessariamente a relação de causa e efeito, dai a limitação do método. Independente da ausência causal entre as 2 variáveis ainda assim a regressão é verdadeira. Pode-se prever o comportamento da Vd sabendo-se a Vi.