Inferência Estatística

Aula 4

Agosto de 2008

Mônica Barros

Conteúdo - Revisão de Probabilidade

Esperança matemática

- **Definição (k-ésimo momento)**
- **O k-ésimo momento de X, ou k-ésimo momento da distribuição de X, denotado por $E(X^k)$ é definido por:**

$$E(X^k)=\begin{cases}\int_{-\infty}^{\infty} x^k.f(x)dx & \text{se X é v.a. contínua}\\ \sum_{\text{todo x}} x^k.f(x)=\sum_{\text{todo x}} x^k.\Pr(X=x) & \text{se X é v.a. discreta}\end{cases}$$

- **Acima k é um inteiro maior ou igual a zero. Em particular, se k = 0, $E(X^0) = E(1) = 1$.**

Esperança matemática

- **Definição (média ou valor esperado)**
- **A média (ou valor esperado ou primeiro momento) de uma variável aleatória é definida como:**

$$\mu=E(X)=\begin{cases}\int_{-\infty}^{\infty} x.f(x)dx & \text{se X é v.a. contínua}\\ \sum_{\text{todo x}} x.f(x)=\sum_{\text{todo x}} x.\Pr(X=x) & \text{se X é v.a. discreta}\end{cases}$$

- **A média de uma variável aleatória é uma medida de tendência central da distribuição de probabilidade desta variável aleatória.**

Esperança matemática

- **Podemos combinar a definição do k-ésimo momento e a definição da média para produzir o *k-ésimo momento central*, como a seguir:**

$$E\left((X-\mu)^k\right)=\begin{cases}\int_{-\infty}^{\infty}(x-\mu)^k.f(x)dx & \text{se X é v.a. contínua}\\ \sum_{\text{todo x}}(x-\mu)^k.f(x)=\sum_{\text{todo x}}(x-\mu)^k.\Pr(X=x) & \text{se X é v.a. discreta}\end{cases}$$

- **Em particular, se k = 1: E(X – μ) = 0, ou seja, o primeiro momento central é sempre nulo.**

Esperança matemática

- **Se k = 1 então:**

$$E(X-\mu)=\int_{-\infty}^{\infty}(x-\mu)f(x)dx=\int_{-\infty}^{\infty}x.f(x)dx-\int_{-\infty}^{\infty}\mu.f(x)dx=$$
$$=\int_{-\infty}^{\infty}x.f(x)dx-\mu\int_{-\infty}^{\infty}f(x)dx=\int_{-\infty}^{\infty}x.f(x)dx-\mu(1)=$$
$$=\mu-\mu=0$$

- **(Idem para o caso discreto)**

Esperança matemática

- **Definição (Variância)**
- **A variância de uma variável aleatória mede a dispersão da distribuição de probabilidade, e é definida como o 2o. momento central:**

$$\sigma^2=VAR(X)=E\left((X-\mu)^2\right)=\begin{cases}\int_{-\infty}^{\infty}(x-\mu)^2.f(x)dx & \text{se X contínua}\\ \sum_{\text{todo x}}(x-\mu)^2.f(x)=\sum_{\text{todo x}}(x-\mu)^2.\Pr(X=x) & \text{se X discreta}\end{cases}$$

Função Geradora de Momentos

$$M(t)=E\left(e^{tX}\right)=\begin{cases}\int_{-\infty}^{\infty}e^{tx}f(x)dx & \text{se X é v.a. contínua}\\ \sum_{\text{todo x}}e^{tx}.f(x)=\sum_{\text{todo x}}e^{tx}.\Pr(X=x) & \text{se X é v.a. discreta}\end{cases}$$

- **A função geradora de momentos é um dos objetos mais úteis em Probabilidade, mas esta função nem sempre existe. Por isso, em cursos mais avançados usa-se a função característica, que sempre existe. O uso da função característica é mais complicado que o da função geradora de momentos, pois envolve números complexos.**
- **Freqüentemente iremos abreviar "função geradora de momentos" por fgm.**

Propriedades da Função Geradora de Momentos

1) A função geradora de momentos é *única* e determina completamente a distribuição da variável aleatória.

- Assim, se duas variáveis aleatórias têm a mesma função geradora de momentos, elas têm a *mesma* distribuição.

Propriedades da Função Geradora de Momentos

- Logo, se você tiver uma tabela de funções geradoras de momentos, você poderá *identificar* qual a densidade que corresponde a uma dada função geradora de momentos.
- Isso é análogo ao uso de uma tabela de transformadas de Laplace ou Fourier (dada a função, sabe-se qual a sua transformada, e vice-versa).

Propriedades da Função Geradora de Momentos

2) M(0) = 1

3) Se M(t) existe para t $\in$ (-h, h) então as derivadas de todas as ordens existem em t = 0 . Além disso, as derivadas de M(t) em t = 0 fornecem os momentos de X.

$$\boxed{M^{(k)}(0) = \left.\frac{d^k M(t)}{dt^k}\right|_{t=0} = E(X^k)}$$

A existência do k-ésimo momento da distribuição implica na existência dos momentos de ordem menor que k. Em particular, a existência da variância implica na existência da média (mas a recíproca não é verdadeira).

Função Geradora de Momentos - exemplo

- Considere um jogo no qual você pode ganhar 0, 1 ou 2 reais, ou perder 2 ou 1 reais com as probabilidades especificadas na tabela a seguir.

x	Pr(X = x)
-2	0.20
-1	0.10
0	0.40
1	0.10
2	0.20

Calcule a função geradora de momentos de X e, a partir dela, encontre a média da distribuição (que representa o seu ganho esperado nesse jogo).

Verifique que a média encontrada por este procedimento é a mesma que a média encontrada pela definição.

Função Geradora de Momentos - exemplo

- Pela definição de média:

$$E(X)=\sum_{x=-2}^{2}x.\Pr(X=x)=\frac{-4}{10}+\left(\frac{-1}{10}\right)+0+\frac{1}{10}+\frac{4}{10}=0$$

- Ou seja, o ganho esperado neste jogo é zero.
- A função geradora de momentos é:

$$M(t)=E\left(e^{tX}\right)=\sum_{x=-2}^{2}e^{tx}.\Pr(X=x)=\left(\frac{2}{10}\right)e^{-2t}+\left(\frac{1}{10}\right)e^{-t}+\left(\frac{4}{10}\right)+\left(\frac{1}{10}\right)e^{t}+\left(\frac{2}{10}\right)e^{2t}$$

- A primeira derivada da função geradora de momentos é:

$$\frac{dM(t)}{dt}=\left(\frac{-4}{10}\right)e^{-2t}+\left(\frac{-1}{10}\right)e^{-t}+\left(\frac{1}{10}\right)e^{t}+\left(\frac{4}{10}\right)e^{2t}$$

Função Geradora de Momentos - exemplo

- A média de X é a primeira derivada da função geradora de momentos em zero, isto é:

$$\left.\frac{dM(t)}{dt}\right|_{t=0}=\left(\frac{-4}{10}\right)+\left(\frac{-1}{10}\right)+\left(\frac{1}{10}\right)+\left(\frac{4}{10}\right)=0$$

- O que concorda com o resultado encontrado antes (pela definição da média).

Função Geradora de Momentos - exemplo

- Nota: Teorema Binomial
- Na demonstração do próximo exemplo empregaremos um resultado conhecido como Teorema Binomial, que é apenas uma maneira de expandir um polinômio.

$$(a+b)^n=\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}$$ onde a e b são número reais e k, n são inteiros ≥ 0

Função Geradora de Momentos - exemplo

- Seja X uma variável aleatória discreta com distribuição Binomial e parâmetros n e p, dada por:

$$\Pr(X=x)=f(x)=\binom{n}{x}.p^x.q^{n-x}$$ onde $q=1-p$ e $x=0,1,2,\ldots n$

- Calcule a função geradora de momentos e, a partir dela, encontre a média e a variância de X.
- Solução
- Pela definição da fgm:

$$M(t)=E(e^{tX})=\sum_{x=0}^{n}\binom{n}{x}.e^{tx}.p^x.q^{n-x}=\sum_{x=0}^{n}\binom{n}{x}.\left(pe^t\right)^x.q^{n-x}=\left(pe^t+q\right)^n=\left(pe^t+1-p\right)^n$$

- Onde, na penúltima igualdade acima, aplicamos diretamente o teorema Binomial.

Função Geradora de Momentos - exemplo

- A primeira derivada da fgm é:

$$M'(t) = n.\left(pe^t + q\right)^{n-1}.\left(pe^t\right)$$

- A média da distribuição é obtida avaliando-se esta derivada em t = 0, isto é:

$$M'(0) = E(X) = n\left(p + q\right)^{n-1}.\left(p\right) = np \longrightarrow$$ **Onde usamos o fato: p + q = 1**

- A segunda derivada de M(t) é:

$$M''(t) = n.p.e^t.\left(q + npe^t\right).\left(q + pe^t\right)^{n-2}$$

Função Geradora de Momentos - exemplo

- Ao avaliarmos esta 2a. derivada em t = 0 encontramos o 2o. momento, $E(X^2)$:

$$M''(0) = E(X^2) = np.(q + np)(q + p)^{n-2} = np.(q + np)$$

- A variância é obtida a partir dos dois primeiros momentos:

- $VAR(X) = E(X^2) - \{E(X)\}^2 = npq + (np)^2 - (np)^2 = npq$

Função Geradora de Momentos - exemplo

- Seja X uma variável aleatória com fgm
- $M(t) = \exp(t^2/2)$
- Podemos obter os momentos de X por diferenciação, mas neste caso há uma maneira mais inteligente de calculá-los, que nos fornece uma expressão geral para todos os momentos de X.

- A expansão de Taylor de $\exp(t^2/2)$ ao redor de zero é:

Função Geradora de Momentos - exemplo

$$e^{t^2/2} = 1 + \frac{t^2}{2} + \frac{1}{2!}.\left(\frac{t^2}{2}\right)^2 + \dots + \frac{1}{k!}.\left(\frac{t^2}{2}\right)^k + \dots =$$

$$= 1 + \frac{t^2}{2} + \frac{t^4}{8} + \dots + \frac{t^{2k}}{(k!)(2^k)} + \dots$$

- Mas, em geral, a expansão de Taylor de uma função geradora de momentos M (t) em torno de zero é:

$$E(e^{tX}) = 1 + t.E(X) + \frac{t^2.E(X^2)}{2!} + \dots + \frac{t^k.E(X^k)}{k!} + \dots$$

Função Geradora de Momentos - exemplo

- Igualando estas duas expressões termo a termo encontramos:

$$E(X^{2k}) = \frac{(2k!)}{(k!)(2^k)} \quad \text{potências pares}$$

$$E(X^{2k-1}) = 0 \quad \text{potências ímpares}$$

- onde k =1, 2, 3, ...
- Veremos que esta é a fgm de uma variável Normal com média 0 e variância 1.

Função Geradora de Momentos – exemplo (para casa)

- Seja X uma variável aleatória discreta com função de probabilidade:

$$f(x) = q.p^{x-1}, \ \text{x = 1,2,3... e q = 1 - p onde} \ 0 < p < 1$$

- Calcule a função geradora de momentos de X.
- A partir da função geradora de momentos, encontre a média e a variância de X.
- *Dica* - use a série geométrica, isto é:

$$\sum_{k=0}^{\infty} a^k = 1 + a + a^2 + a^3 + ... = \frac{1}{1-a} \quad \text{desde que } |a| < 1$$

Função Geradora de Momentos – exemplo (para casa)

- Seja X uma variável aleatória com função de probabilidade Poisson com média λ.
- Encontre a função geradora de momentos e use-a para mostrar que:
  - E(X) = VAR(X) = λ
- Dica: lembre-se da série de Taylor da exponencial, isto é:

$$e^u = \sum_{k=0}^{\infty} \frac{u^k}{k!} = 1 + u + \frac{u^2}{2!} + \frac{u^3}{3!} + \frac{u^4}{4!} + ...$$

A soma de duas variáveis aleatórias independentes

- *Este é um dos problemas mais importantes em Probabilidade. Por que?*

- Freqüentemente iremos encontrar conjuntos de variáveis aleatórias que são independentes e identicamente distribuídas (isto é, variáveis que, num certo sentido, são “todas iguais” pois têm a mesma distribuição de probabilidade e uma não afeta as outras).

A soma de duas variáveis aleatórias independentes

- Este conjunto de variáveis é chamado de variáveis iid, ou de amostra aleatória.
- Suponha que a média de todos os X_i ´s é desconhecida.
- Um estimador (ou "chute") que parece bastante sensato para a média da distribuição dos X_i´s é:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{a média amostral}$$

- Mas, para calcularmos a média amostral precisamos saber qual a distribuição da soma dos X_i´s, o que é um problema ainda mais complicado do que o que estamos propondo agora!

A soma de duas variáveis aleatórias independentes

- Teorema (Fórmula da Convolução)
- Seja $Y = X_1+X_2$ onde X_1 e X_2 são variáveis independentes com densidade conjunta $f(x_1, x_2) = f_1(x_1).f_2(x_2)$.
- Então a densidade de Y é dada por:
- No caso contínuo

$$g(y) = \int_{-\infty}^{\infty} f_1(x_1).f_2(y-x_1)dx_1 = \int_{-\infty}^{\infty} f_1(y-x_2).f_2(x_2)dx_2$$

- No caso discreto

$$g(y) = \sum_{\text{todo } x_1} f_1(x_1).f_2(y-x_1) = \sum_{\text{todo } x_2} f_1(y-x_2).f_2(x_2)$$

Exemplo

- Suponha que X e Y são independentes e identicamente distribuídas com densidade Exponencial(λ).
- Seja Z = X + Y. Encontre a densidade de Z.
- Solução
- A densidade de X é: $\lambda.e^{-\lambda x}$, x > 0.
- A densidade de Y é : $\lambda.e^{-\lambda y}$, y > 0.
- Como X e Y são independentes, a densidade conjunta é apenas o produto destas marginais.
- A densidade de Z = X + Y é obtida através da fórmula da convolução.

Exemplo

$$g(z) = \int_0^{\infty} f_x(x).f_y(z-x)dx = \int_0^{\infty} \lambda.e^{-\lambda x}.\lambda\left(e^{-\lambda(z-x)}\right)dx$$

- Mas, o integrando acima é positivo se, e somente se, ambos os termos são positivos. Para que isto aconteça é necessário que x > 0 e, ao mesmo tempo, z - x > 0, isto é, x < z. Logo, os limites de integração são 0 e z na integral anterior.

$$g(z) = \int_0^{z} \lambda.e^{-\lambda x}.\lambda e^{-\lambda z}.e^{+\lambda x}dx = \lambda^2.e^{-\lambda z}\int_0^{z} dx = \lambda^2.e^{-\lambda z}\left(x\right)\Big|_0^z = \lambda^2.z.e^{-\lambda z}$$

- Ou seja, a soma de duas variáveis Exponenciais independentes NÃO É uma variável Exponencial.

Exemplo

- Sejam X e Y variáveis independentes discretas com distribuição Poisson com média k. Encontre a função de probabilidade de Z=X+Y.

$$f(x) = \Pr(X = x) = \frac{e^{-k}.k^x}{x!} \text{ onde } x = 0,1,2,......$$

e

$$f(y) = \Pr(Y = y) = \frac{e^{-k}.k^y}{y!} \text{ onde } y = 0,1,2,......$$

Exemplo

- Pela fórmula da convolução:

$$g(z) = \sum_{x=0}^{\infty} \frac{e^{-k}.k^x}{x!}.\frac{e^{-k}.k^{z-x}}{(z-x)!} = e^{-2k}\sum_{x=0}^{z}\frac{1}{x!(z-x)!}.k^x.k^{z-x} = \frac{e^{-2k}}{z!}\sum_{x=0}^{z}\frac{z!}{x!(z-x)!}.k^x.k^{z-x} =$$

$$= \frac{e^{-2k}}{z!}\sum_{x=0}^{z}\binom{z}{x}.k^x.k^{z-x} = \frac{e^{-2k}}{z!}.(k+k)^z = \frac{e^{-2k}.(2k)^z}{z!}$$

- Note que o somatório (que ia de x = 0 a ∞) passou a ser um somatório de x = 0 até x = z. Isto acontece por que precisamos garantir que ambas as funções de probabilidade no somatório sejam maiores que zero, o que só ocorre se AMBOS x > 0 e z-x > 0. Na penúltima igualdade à direita usamos o teorema Binomial para concluir que o somatório era igual a $(2k)^z$.

Exemplo

- O que existe de interessante neste exemplo?
- Note que, ao contrário do exemplo anterior, aqui "preservamos" a mesma "família" de funções de probabilidade.

- Em outras palavras: ***a soma de duas Poisson independentes é Poisson***, mas a soma de duas Exponenciais independentes não é Exponencial!

Exemplo

- Sejam X e Y v.a. independentes com função de probabilidade Geométrica(p). Encontre a função de probabilidade de Z = X+Y.
- Solução
- Aqui novamente usamos a fórmula da convolução.

$$g(z) = \sum_{x=1}^{\infty} p.(1-p)^{x-1}.p.(1-p)^{(z-x)-1} =$$

$$= \sum_{x=1}^{z-1} p^2.(1-p)^{z-2} = p^2.(1-p)^{z-2}.\sum_{x=1}^{z-1} 1 = (z-1)\,p^2.(1-p)^{z-2}$$

Exemplo

- No último passo acima o somatório não envolve x e existem (z-1) termos no somatório.

- Note que os limites do somatório são 1 e z-1 pois precisamos satisfazer simultaneamente às condições x > 1 e z- x > 1.

- A função de probabilidade condicional de X dado Z é obtida pela razão entre a conjunta de X e Z e a marginal de Z, recém calculada. Mas, a conjunta de X e Z é a conjunta de X e X + Y, ou seja, é a mesma coisa que a conjunta de X e Y.

Exemplo

- Esta função de probabilidade conjunta é, pela independência de X e Y, apenas o produto das marginais. Logo a conjunta de X e Z é:

$$f(x,z) = p.(1-p)^{x-1}.p(1-p)^{(z-x)-1} = p^2(1-p)^{z-2}$$

- A função de probabilidade condicional de X dado X+Y = z é então:

$$g(x \mid Z = z) = \frac{f(x,z)}{g(z)} = \frac{p^2.(1-p)^{z-2}}{(z-1).p^2.(1-p)^{z-2}} = \frac{1}{z-1}$$

onde x = 1, 2, ... z - 1

- Ou seja, X dado a soma é Uniforme Discreta nos inteiros 1, 2, .., z-1.

Para Casa (problemas de prova)

- 1) Considere o Exemplo da soma de Poissons
- Encontre a distribuição condicional de X dado Z = X + Y.

- 2) Sejam X e Y independentes e identicamente distribuídos Bin(n, p). Mostre que Z = X + Y é Binomial (2n, p) e encontre a distribuição condicional de X dado Z = X + Y.

Combinações Lineares de v.a. Independentes

- Objetivos
  - Dados $X_1, X_2, ..., X_n$ independentes (mas não necessariamente identicamente distribuídos) tais que: $E(X_i) = \mu_i$ e $VAR(X_i) = \sigma_i^2$, encontrar a média, variância e função geradora de momentos de:

$$Y = a_0 + \sum_{i=1}^{n} a_i X_i$$

  - em função dos momentos dos X_i´s. Note que, na definição de Y, os a_i's são constantes.
  - Também, a independência dos X_i´s assegura que todas as covariâncias (e correlações) são nulas.

Combinações Lineares de v.a. Independentes

- **A média de Y é:**

$$E(Y) = E\left(a_0 + \sum_{i=1}^{n} a_i.X_i\right) = a_0 + \sum_{i=1}^{n} a_i.E(X_i) =$$
$$= a_0 + \sum_{i=1}^{n} a_i.\mu_i \quad \text{onde } \mu_i = E(X_i)$$

- **A variância de Y é:**

$$VAR(Y) = \sum_{i=1}^{n} a_i^2.VAR(X_i) = \sum_{i=1}^{n} a_i^2.\sigma_i^2$$
onde $VAR(X_i) = \sigma_i^2$

- **Note que a constante a_0 não afeta a variância, e os outros a_i´s aparecem ao quadrado (por que?).**

Combinações Lineares de v.a. Independentes

- **A função geradora de momentos de Y é:**

$$M_Y(t) = E\left(e^{tY}\right) = E\left(e^{t(a_0 + a_1X_1 + \ldots + a_nX_n)}\right)$$
$$= E\left(e^{ta_0}e^{ta_1X_1}....e^{ta_nX_n}\right) =$$
$$= e^{ta_0}E\left(e^{ta_1X_1}....e^{ta_nX_n}\right)$$
e como conseqüência da independência
$$= e^{ta_0}E\left(e^{ta_1X_1}\right)E\left(e^{ta_2X_2}\right)...E\left(e^{ta_nX_n}\right) =$$
$$e^{ta_0}M_{X_1}(ta_1)M_{X_2}(ta_2)...M_{X_n}(ta_n)$$

Combinações Lineares de v.a. Independentes

- **Um caso particular importante (e que veremos muitas vezes daqui em diante) é o da soma de variáveis independentes.**

- **O ponto importante aqui é: se pudermos associar a fgm de Y a alguma fgm conhecida, então automaticamente sabemos qual é a densidade (ou função de probabilidade) de Y, pela unicidade da fgm.**

A soma de v.a. independentes

- **Sejam $X_1, X_2, ..., X_n$ v.a. independentes mas não necessariamente identicamente distribuídas.**

- **Seja:** $Y = \sum_{i=1}^{n} X_i$
- **Então:**

$$E(Y) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \mu_i \quad \text{onde } \mu_i = E(X_i)$$

A soma de v.a. independentes

- A variância de Y é:

$$VAR(Y) = \sum_{i=1}^{n} VAR(X_i) = \sum_{i=1}^{n} \sigma_i^2 \quad \text{onde} \quad VAR(X_i) = \sigma_i^2$$

- A fgm de Y é:

$$M_Y(t) = E\left(e^{tY}\right) = E\left(e^{t(X_1+\ldots+X_n)}\right) = E\left(e^{tX_1} \ldots e^{tX_n}\right) =$$

e como conseqüência da independência

$$= E\left(e^{tX_1}\right)E\left(e^{tX_2}\right)\ldots E\left(e^{tX_n}\right) = M_{X_1}(t)M_{X_2}(t)\ldots M_{X_n}(t)$$

- Ou seja, a expressão da fgm de Y é muito simples, apenas o produto das fgms das variáveis independentes que compõem a soma!

A Média Amostral

- Sejam $X_1, X_2, \ldots, X_n$ v.a. independentes e identicamente distribuídas (iid).
- Em particular, o fato dos X´s serem iid implica em $E(X_i) = \mu$, $VAR(X_i) = \sigma^2$ e $COV(X_i, X_j) = 0$ para todo par i,j.
- Considere a média amostral calculada a partir destes X_i´s, isto é:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

A Média Amostral

- Para que serve a média amostral?
- Suponha que existe esta coleção de X_i´s, todos eles iid, ou seja, todos eles provêm da mesma densidade ou função de probabilidade.
- Suponha que a média desta densidade é μ, um número desconhecido.
- A média amostral é um estimador "sensato" para a média da distribuição (μ), quando esta última for desconhecida.

A Média Amostral

- Podemos encontrar a média, variância e fgm da média amostral como casos particulares dos resultados para combinações lineares de variáveis independentes.
- É fácil mostrar que:

$$E\left(\overline{X}\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{n\mu}{n} = \mu$$

(A mesma que a média de qualquer X_i)

A Média Amostral

- Mas, a variância da média amostral é:

$$VAR(\overline{X}) = \sum_{i=1}^{n} \frac{1}{n^2} VAR(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}$$

(A variância de $\overline{X}$ é menor que a de X_i, desde que usemos mais de uma observação !)

- **Moral da história: é vantajoso usar a média amostral para "chutar" a média da densidade dos X´s, pois ela implica numa redução da variância se comparada ao uso de um X_i sozinho.**

A Média Amostral

- **Estes resultados justificam o uso da média amostral como um estimador "sensato" de μ.**
- **Note que, à medida que tomamos uma amostra cada vez maior (n cresce), a variância da média amostral decresce, ou seja, a "dispersão" do estimador decresce. Além disso, a média de $\overline{X}$ é μ, que é a quantidade que ele pretende estimar.**

A Média Amostral

- **A fgm da média amostral também pode ser encontrada como um caso particular da combinação linear de v.a. independentes.**
- **Seja Y a soma de $X_1, X_2, ..., X_n$, que são todos iid. Então:**

$$M_{\overline{X}}(t) = E(e^{t\overline{X}}) = E(e^{tY/n}) = M_Y(t/n) = \left[M_{X_1}(t/n)\right]^n$$

Combinações Lineares de v.a. Dependentes

- **Este é um caso mais geral (e mais complicado que o anterior).**
- **Suponha que $E(X_i) = \mu_i$, $VAR(X_i) = \sigma_i^2$, e $COV(X_i, X_j) = \rho_{ij}.\sigma_i.\sigma_j \neq 0$ para algum par i,j.**
- **Queremos encontrar a média, variância e fgm de Y definido como:**

$$Y = a_0 + \sum_{i=1}^{n} a_i X_i$$

Combinações Lineares de v.a. Dependentes

- A expressão da média é a mesma que no caso anterior (variáveis independentes), isto é:

$$E(Y) = E\left(a_0 + \sum_{i=1}^{n} a_i.X_i\right) = a_0 + \sum_{i=1}^{n} a_i.E(X_i) =$$

$$= a_0 + \sum_{i=1}^{n} a_i.\mu_i \quad \text{onde } \mu_i = E(X_i)$$

- A expressão para a variância torna-se mais complicada, pois deverá levar em conta as interdependências entre as variáveis. Pode-se mostrar que:

Combinações Lineares de v.a. Dependentes

$$VAR(Y) = \sum_{i=1}^{n} a_i^2.VAR(X_i) + \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j COV\left(X_i, X_j\right)$$

$$= \sum_{i=1}^{n} a_i^2.\sigma_i^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \rho_{ij}\sigma_i\sigma_j \quad \text{onde no 2o. termo i } \neq \text{ j}$$

onde ρ_{ij} é o coeficiente de correlação entre X_i e X_j

Ou seja, a dependência entre os X´s leva à presença de um termo adicional no cálculo da variância, devido às covariâncias entre os diversos X´s.

Combinações Lineares de v.a. Dependentes

- A expressão anterior, que nos dá a variância de uma combinação linear de variáveis dependentes, é essencial para medir o risco de um portfolio ou carteira, como veremos mais tarde.

- Na verdade, a dependência entre ativos financeiros (que são variáveis aleatórias) pode ser usada de maneira proveitosa para reduzir o risco de um portfolio, num processo que chamamos de “diversificação”.

Combinações Lineares de v.a. Dependentes

- A função geradora de momentos de Y não pode, em geral, ser simplificada como no caso anterior (em que os X´s eram independentes).

Portfolios (Carteiras)

- Considere uma coleção de n ativos financeiros. Um portfolio (ou carteira) é apenas uma combinação linear destes ativos na qual a soma dos pesos é 1 (ou 100%).
- Algumas vezes requer-se também que todos os pesos sejam positivos, mas isso não é estritamente necessário.

Retorno Aritmético

- Suponha que P_t seja o preço de um ativo no instante t. O retorno (aritmético) deste ativo entre os instantes t-1 e t é:

$$RA_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

- Em Finanças "quantitativas", o retorno aritmético é pouco empregado, pois possui um grave inconveniente: como o preço de um ativo não pode ser menor que zero, o retorno aritmético tem um valor mínimo (quando $P_t = 0$), igual a –1.

Retorno Geométrico

- Em Finanças, os retornos serão tratados como variáveis aleatórias, e um dos modelos mais simples supõe que os retornos são Normais (que iremos estudar no capítulo 7 do meu livro de Probabilidade). Mas, uma v.a. Normal é ilimitada. Então, o que fazer?
- Trabalhar com o retorno Geométrico, definido como:

$$R_t = \log\left(\frac{P_t}{P_{t-1}}\right) \quad \text{onde } \log \text{ denota o logaritmo na base "e"}$$

Retorno Geométrico

- O retorno geométrico é definido em toda a reta, então pode-se modelá-lo pela distribuição Normal.
- Retornos Geométricos são aditivos. Por exemplo, se queremos encontrar o retorno entre os dias 1 e 4:

$$R_{1,4} = \log\left(\frac{P_4}{P_1}\right) = \log\left(\frac{P_4}{P_3}\frac{P_3}{P_2}\frac{P_2}{P_1}\right) = \log\left(\frac{P_4}{P_3}\right) + \log\left(\frac{P_3}{P_2}\right) + \log\left(\frac{P_2}{P_1}\right)$$

- Ou seja, o retorno Geométrico entre os dias 1 e 4 é apenas a soma dos retornos geométricos diários no período.
- Se a variação do período é pequena, os retornos geométrico e aritmético são aproximadamente iguais.

Risco e Retorno de um Portfolio

- Sejam $X_1, X_2, ..., X_n$ os retornos de n ativos, tais que $E(X_i) = \mu_i$, $VAR(X_i) = \sigma_i^2$, e $COV(X_i, X_j) = \rho_{ij}.\sigma_i.\sigma_j \neq 0$.
- Seja P uma carteira, isto é:

$$P = \sum_{i=1}^{n} w_i X_i \quad \text{onde} \quad \sum_{i=1}^{n} w_i = 1$$

- Qual o retorno esperado da carteira, e qual o seu risco (medido pela sua variância ou desvio padrão)? Tudo isso pode ser encontrado a partir dos resultados mostrados nesta aula.

Risco e Retorno de um Portfolio

- O retorno esperado (ou retorno médio) do portfolio é:

$$E(P) = \sum_{i=1}^{n} w_i E(X_i) = \sum_{i=1}^{n} w_i \mu_i$$

- O risco do portfolio é dado pela sua variância (ou, preferencialmente, pelo seu desvio padrão, que está nas mesmas unidades que o retorno).

Risco e Retorno de um Portfolio

$$\boxed{VAR(P) = \sum_{i=1}^{n} w_i^2 .\sigma_i^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \rho_{ij}\sigma_i\sigma_j}$$

onde ρ_{ij} é o coeficiente de correlação entre X_i e X_j

- Qual o efeito do segundo termo na equação acima? Imagine que conseguíssemos encontrar ativos com correlações negativas. O efeito destes termos seria a redução da variância do portfolio! Isto é o que está por trás da idéia de *diversificação*. A gente pode reduzir o risco sem, necessariamente, diminuir o retorno.

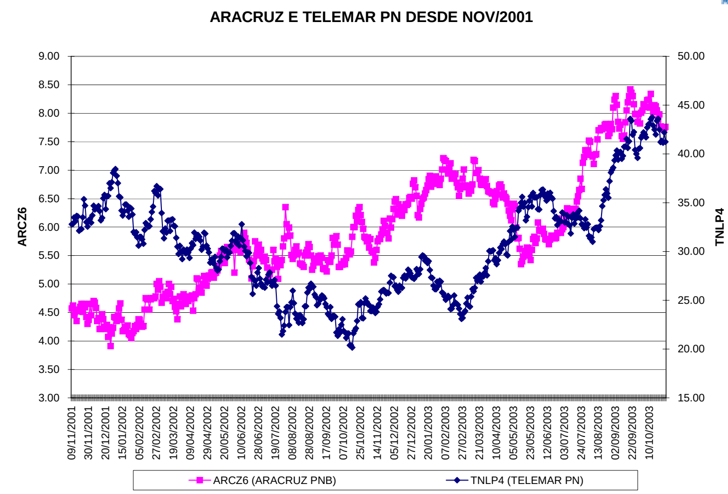
Risco e Retorno de um Portfolio

- A otimização de um portfolio é feita escolhendo-se os pesos w_i's tal que alguma restrição é satisfeita.
- Um caso típico é: encontre os pesos tal que, para um retorno especificado, o risco é mínimo.
- Pode-se reescrever este problema como:
  - Dado E(P) = constante, encontre os pesos w_i (i= 1,2,...n) sujeitos à restrição $\sum_{i=1}^{n} w_i = 1$
  - e que minimizem VAR(P).

Exemplo

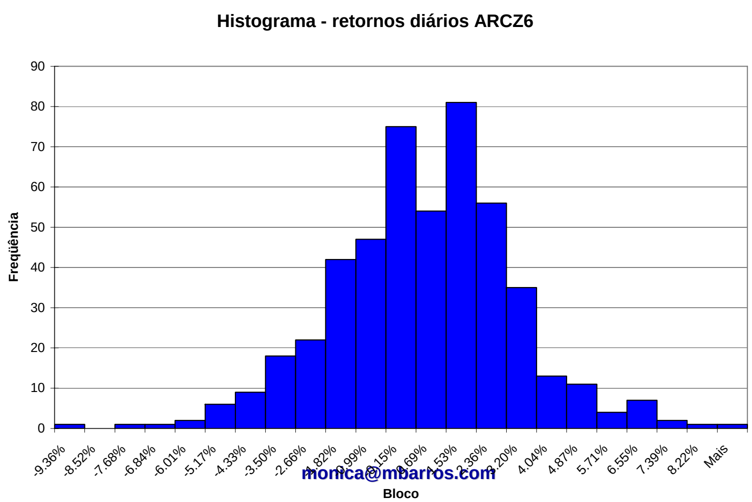
- A seguir trabalhamos com retornos diários e construímos portfolios compostos pelas seguintes ações:
  - TNLP4 (Telemar PN)
  - ARCZ6 (Aracruz Celulose PN)
- Os retornos diários (geométricos) foram calculados entre 12/11/2001 e 29/10/2003.
- O gráfico a seguir apresenta a evolução dos preços das duas ações no período.

Exemplo



Exemplo

- O histograma dos retornos diários de ARCZ6 é:



Exemplo

- A correlação entre os retornos de TNLP4 e ARCZ6 é 0.1153, e a média e o desvio padrão dos retornos de cada uma das séries é:

	Retornos	
	ARCZ6	TNLP4
média	0.11%	0.05%
d,p.	2.48%	2.52%

- Seja P = α.ARCZ6 + (1-α)TNLP4 o portfolio formado por α% de Aracruz e (1-α)% de Telemar. Como variam o retorno esperado e o risco do portfolio à medida que alteramos α?

Exemplo

α (% de ARCZ6 no portfolio)	retorno esperado	risco do portfolio(desvio padrão)
0.00	0.047%	2.52%
0.05	0.050%	2.41%
0.10	0.053%	2.31%
0.15	0.056%	2.22%
0.20	0.059%	2.13%
0.25	0.062%	2.06%
0.30	0.065%	1.99%
0.35	0.068%	1.94%
0.40	0.071%	1.90%
0.45	0.074%	1.88%
0.50	0.078%	1.87%
0.55	0.081%	1.87%
0.60	0.084%	1.89%
0.65	0.087%	1.92%
0.70	0.090%	1.97%
0.75	0.093%	2.03%
0.80	0.096%	2.10%
0.85	0.099%	2.18%
0.90	0.102%	2.27%
0.95	0.105%	2.37%
1.00	0.108%	2.48%

Estes são o retorno e o risco do portfolio computados ao longo de todo o período (isto é, usando todos os retornos diários disponíveis)

Na próxima página exibimos este gráfico em termos de α.

Qual o α que leva ao portfolio de menor risco? Você consegue enxergar que, no caso de um portfolio com 2 ativos, a solução para este problema pode ser obtida analiticamente, e de maneira bem simples?

Exemplo