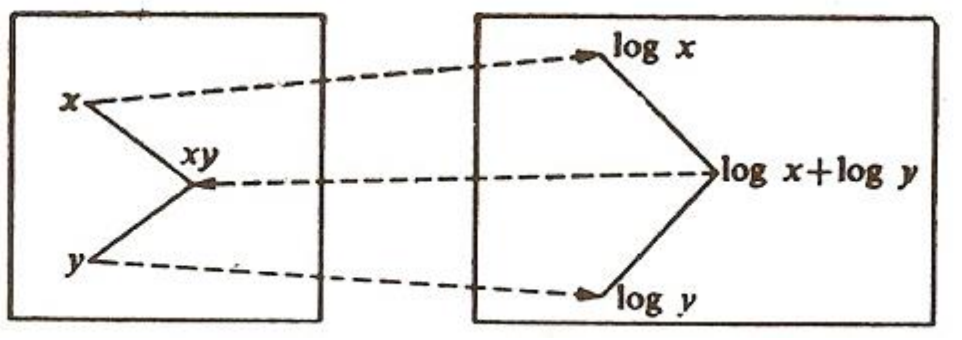
10 A Função Geratriz de Momentos

ESQUEMA DO CAPÍTULO

10.1 Introdução

- **Transformações matemáticas:**

  Em lugar de executar certas contas diretamente com os números x e y, pode ser mais simples transformá-los por meio de alguma função, fazer as contas e depois transformá-los novamente, em um esquema conforme ilustrado na figura.



10.2 A Função Geratriz de Momentos

- **Definição:**

  Seja X uma variável aleatória discreta, com distribuição de probabilidade $p(x_i) = P(X = x_i)$, $i = 1, 2, \ldots$. A função M_X, denominada ***função geratriz de momentos*** de X, é definida por

$$M_X(t) = \sum e^{tx_j} p(x_j).$$

  Se X for uma variável aleatória contínua com fdp f, definiremos a função geratriz de momentos por

$$M_X(t) = \int_{-\infty}^{+\infty} e^{tx} f(x)dx.$$

- **Comentários:**

a) Em qualquer dos casos, o discreto ou o contínuo, $M_X(t)$ é apenas o ***valor esperado*** de e^{tX}. Por isso, podemos combinar as expressões acima e escrever

$$M_X(t) = E(e^{tX}).$$

10.2 A Função Geratriz de Momentos

- **Comentários (cont.):**

b) $M_X(t)$ é o valor que a função M_X toma para a variável (real) *t*. A notação, que indica a dependência de X, é empregada porque nós poderemos desejar considerar duas variáveis aleatórias, X e Y, e depois investigarmos a função geratriz de momentos de cada uma delas, a saber M_X e M_Y.

c) É normalmente empregada a abreviatura fgm para designá-la.

d) A fgm é escrita como uma série infinita ou integral (imprópria), conforme a variável aleatória seja discreta ou contínua. Essa série (ou integral) poderá nem sempre existir (isto é, convergir para um valor finito) para todos os valores de *t*. Consequentemente, poderá acontecer que a fgm não seja definida para todos os valores de *t*. No entanto, não nos preocuparemos com esta dificuldade potencial e admitiremos que ela ***sempre*** exista (note que a fgm existe sempre para $t = 0$ e é igual a 1).

10.2 A Função Geratriz de Momentos

- **Comentários (final):**

e) Existe uma outra função, estreitamente relacionada com a fgm, a qual é frequentemente empregada em seu lugar.

Ela é denominada ***função característica***, denotada por C_X e definida por

$$C_X(t) = E(e^{itX}),$$

em que i é a unidade imaginária, $i = \sqrt{-1}$.

Por motivos teóricos, existe enorme vantagem em empregar-se $C_X(t)$ em lugar de $M_X(t)$

[Por esta razão: $C_X(t)$ sempre existe, para todos os valores de t].

Contudo, a fim de evitarmos cálculos com números complexos, restringiremos nossa exposição à função geratriz de momentos.

f) Adiaremos até a Sec. 10.4 a justificativa para denominar M_X de função geratriz de momentos.

10.3 Exemplos de Funções Geratrizes de Momentos

- **Exemplo:**

  Suponha-se que X seja ***uniformemente distribuída*** sobre o intervalo [a, b]. Logo, a fgm será dada por

$$\mathrm{M}_{\mathrm{X}}(t) = \int_a^b \frac{e^{tx}}{b-a}\,dx = \frac{1}{b-a}\left[e^{bt} - e^{at}\right],\ t \neq 0.$$

- **Exemplo II:**

  Suponha-se que X seja ***binomialmente distribuída***, com parâmetros n e p. Nesse caso

$$\mathrm{M}_{\mathrm{X}}(t) = \sum_{k=0}^{n} e^{tk}\binom{n}{k} p^k (1-p)^{n-k} = [pe^t + (1-p)]^n.$$

10.4 Propriedades da Função Geratrizes de Momentos

- **Teorema:**

  $M_X^{(n)}(0) = E(X^n).$

  Ou seja, a derivada *n*-ésima de $M_X(t)$, avaliada no ponto $t = 0$, fornece $E(X^n)$.

- **Comentários:**

a) Os números $E(X^n)$, $n = 1, 2, ...$, são denominados os ***momentos de ordem*** *n* da variável aleatória X, em relação a zero. A partir do conhecimento da função M_X, os momentos podem ser gerados. Daí, o nome "função geratriz de momentos".

b) Em particular

$$V(X) = E(X^2) - [E(X)]^2 = M''(0) - [M'(0)]^2.$$

c) Pode parecer surpreendente, mas em muito problemas é mais fácil obter a fgm da variável aleatória X e depois derivá-la, para obter os seus momentos.

10.4 Propriedades da Função Geratrizes de Momentos

- **Exemplo:**

  Suponha-se que X tenha uma distribuição binomial com parâmetros *n* e *p*. Consequentemente (de exemplo anterior)

  $$M_X(t) = [pe^t + q]^n.$$

  Por isso,

  $$M_X'(t) = n(pe^t + q)^{n-1}pe^t,$$

  $$M_X''(t) = np[e^t(n-1)(pe^t + q)^{n-2}pe^t + (pe^t + q)^{n-1}e^t].$$

  Logo,

  $$E(X) = M_X'(0) = np,$$

  o que concorda com resultado anterior.

10.4 Propriedades da Função Geratrizes de Momentos

- **Exemplo (final):**

  Também,

  $$E(X^2) = M_X''(0) = np[(n\text{-}1)p + 1].$$

  Daí,

  $$V(X) = M_X''(0) - [M'(0)]^2 = np(1\text{-}p),$$

  também de acordo com resultado anterior.

10.4 Propriedades da Função Geratrizes de Momentos

- **Teorema II:**

  Suponha-se que a variável aleatória X tenha uma fgm M_X. Seja $Y = \alpha X + \beta$. Então, M_Y, a fgm da variável aleatória Y, será dada por

  $$M_Y(t) = e^{\beta t} M_X(\alpha t).$$

- **Explicação:**

  Para achar a fgm de $Y = \alpha X + \beta$, calcule-se a fgm de X para αt (em vez de t) e multiplique-se por $e^{\beta t}$.

- **Demonstração:**

  $$M_Y(t) = E(e^{Yt}) = E[e^{(\alpha X + \beta)t}] = e^{\beta t}E(e^{\alpha t X}) = e^{\beta t} M_X(\alpha t).$$

10.4 Propriedades da Função Geratrizes de Momentos

- **Teorema III:**

  Sejam X e Y duas variáveis aleatórias com fgm $M_X(t)$ e $M_Y(t)$, respectivamente. Se $M_X(t) = M_Y(t)$ para todos os valores de t, então X e Y terão a mesma distribuição de probabilidade.

- **Comentário:**

  A demonstração deste teorema é bastante difícil para ser dada aqui, mas o importante é que compreenda exatamente o que ele significa, ou seja, a fgm determina univocamente a distribuição de probabilidade da variável aleatória.

10.4 Propriedades da Função Geratrizes de Momentos

- **Exemplo:**

  Suponhamos que X tenha uma distribuição $N(\mu, \sigma^2)$. Seja $Y = \alpha X + \beta$. Então, Y será também distribuída normalmente, com $N(\alpha\mu+\beta, \alpha^2\sigma^2)$.

  De fato, de teorema anteriormente visto, teremos que a fgm de Y será

  $$M_Y(t) = e^{\beta t} M_X(\alpha t).$$

  Também, de exemplo anterior, teremos que:

  $$M_X(t) = e^{(\mu)t + \frac{(\sigma^2)t^2}{2}}.$$

  Portanto,

  $$M_Y(t) = e^{\beta t} M_X(\alpha t) = e^{\beta t}\left[e^{\alpha\mu t + \frac{(\alpha\sigma)^2 t^2}{2}}\right] = e^{(\alpha\mu+\beta)t + \frac{(\alpha\sigma)^2 t^2}{2}}.$$

10.4 Propriedades da Função Geratrizes de Momentos

- **Teorema IV:**

  Suponha-se que X e Y sejam variáveis aleatórias independentes. Façamos $Z = X + Y$. Sejam $M_X(t)$, $M_Y(t)$ e $M_Z(t)$ as fgm das variáveis aleatórias X, Y e Z, respectivamente. Então

  $$M_Z(t) = M_X(t) \times M_Y(t).$$

- **Demonstração:**

  $$M_Z(t) = E(e^{Zt}) = E[e^{(X+Y)t}] = E(e^{Xt+Yt})$$
  $$= E(e^{Xt}) \times E(e^{Yt}) = M_X(t) \times M_Y(t).$$

10.4 Propriedades da Função Geratrizes de Momentos

- **Comentário:**

  Este teorema pode ser generalizado assim. Se $X_1,\ldots, X_n$ forem variáveis aleatórias independentes, com fgm $M_{X_i}(t)$, $i = 1,\ldots, n$, então $M_Z(t)$, a fgm de

  $$Z = X_1 + \ldots + X_n,$$

  será dada por

  $$M_Z(t) = M_{X_1}(t) \times \ldots \times M_{X_n}(t).$$

10.5 Propriedades Reprodutivas

- **Definição:**

  Se duas (ou mais) variáveis aleatórias tiverem uma determinada distribuição, forem somadas e a variável aleatória resultante desta soma tiver uma distribuição do mesmo tipo que aquela das parcelas, dizemos que elas obedecem à ***propriedade reprodutiva*** (ou ***aditiva***).

- **Exemplo:**

  Suponha-se que X e Y sejam variáveis aleatórias independentes, com distribuições $N(\mu_1, \sigma^2{}_1)$ e $N(\mu_2, \sigma^2{}_2)$, respectivamente. Façamos $Z = X + Y$. Consequentemente,

$$\begin{aligned} M_Z(t) &= M_X(t) \times M_Y(t) \\ &= \exp(\mu_1 t + \sigma^2{}_1 t^2/2) \times \exp(\mu_2 t + \sigma^2{}_2 t^2/2) \\ &= \exp[(\mu_1 + \mu_2)t + (\sigma^2{}_1 + \sigma^2{}_2)t^2/2], \end{aligned}$$

  que representa a fgm de uma variável aleatória normalmente distribuída, com valor esperado $\mu_1 + \mu_2$ e variância $\sigma^2{}_1 + \sigma^2{}_2$.

10.5 Propriedades Reprodutivas

- **Teorema (Propriedade aditiva da distribuição normal):**

  Sejam $X_1, ..., X_n$ *n* variáveis aleatórias independentes com distribuição $N(\mu_i; \sigma^2_i)$, $i = 1, ..., n$.

  Façamos $Z = X_1 + ... + X_n$. Então, Z terá distribuição

$$N\left(\sum_{i=1}^{n} \mu_i; \sum_{i=1}^{n} \sigma^2_i \right).$$

- **Demonstração:**

  Deixada como exercício.

10.5 Propriedades Reprodutivas

- **Teorema II:**

  Sejam $X_1,\ldots, X_n$ variáveis aleatórias independentes com distribuição de Poisson, com parâmetro α_i, $i = 1,\ldots, n$. Então $Z = X_1 +\ldots+ X_n$ terá distribuição de Poisson com parâmetro

  $$\alpha = \alpha_1 +\ldots+ \alpha_n.$$

- **Demonstração:**

  Considera-se, inicialmente o caso de $n = 2$, e completa-se por indução matemática

  $$\begin{aligned} M_Z(t) &= M_{X_1}(t) \times M_{X_2}(t) \\ &= e^{\alpha_1(e^t-1)} \times e^{\alpha_2(e^t-1)} \\ &= e^{(\alpha_1 + \alpha_2)(e^t-1)}. \end{aligned}$$

10.5 Propriedades Reprodutivas

- **Teorema III:**

  Suponha-se que a distribuição de X_i seja χ^2_{ni}, $i = 1,\ldots, k$, em que os X_i sejam variáveis aleatórias independentes.

  Então, $Z = X_1 +\ldots+ X_k$ terá distribuição χ^2_n, em que $n = n_1 + \ldots + n_k$.

- **Demonstração:**

  Demonstra-se, pelo produto das fgm da distribuição chi-quadrado

  $$M_{Xi}(t) = (1 - 2t)^{-ni/2},\ i = 1,\ldots, k,$$

  o que fornece

  $$M_Z(t) = (1 - 2t)^{-(n1 + \ldots + nk)/2}.$$

10.5 Propriedades Reprodutivas

- **Teorema IV:**

  Suponha-se que $X_1, \ldots, X_k$ sejam variáveis aleatórias independentes, cada uma com distribuição N(0,1). Então, $S = X^2_1 + \ldots + X^2_k$ terá distribuição χ^2_k.

- **Comentários:**

a) Para $T = \sqrt{S}$, quando $n = 2$, T segue uma distribuição conhecida como ***distribuição de Rayleigh***.

b) Para $T = \sqrt{S}$, quando $n = 3$, T segue uma distribuição conhecida como ***distribuição de Maxwell***.

10.5 Propriedades Reprodutivas

- **Teorema V:**

  Seja $Z = X_1 + \ldots + X_r$, em que os X_i são r variáveis aleatórias independentes, identicamente distribuídas, cada uma tendo distribuição exponencial, com o mesmo parâmetro α. Então Z terá uma distribuição gama, com parâmetros α e r.

- **Comentários:**

a) O teorema não será verdadeiro se os parâmetros das várias distribuições exponenciais forem diferentes, o que fica evidente quando consideramos a fgm da soma das variáveis aleatórias.

b) O seguinte corolário do teorema apresenta considerável importância em algumas aplicações estatísticas. A variável aleatória $W = 2\alpha Z$ tem distribuição χ^2_{2r}. Assim, podemos calcular determinadas probabilidades associadas a Z pelo uso de tábuas da distribuição qui-quadrado, se α e r forem dados. Por exemplo

$$P(Z \leq 3) = P(2\alpha Z \leq 2\alpha 3) = P(W \leq 6\alpha).$$

10.6 Sequencias de Variáveis Aleatórias

- Suponha-se que temos uma sequência de variáveis aleatórias X_1, ..., X_i, ... Cada uma dessas variáveis aleatórias poder ser descrita em termos de F_i, sua fd, em que $F_i(t) = P(X_i \leq t)$, $i = 1, 2, \ldots$. É frequente estarmos interessados no que ocorre a F_i, quando $i \to \infty$, isto é, na ***função distribuição limite*** F correspondente a alguma variável aleatória X, tal que em algum sentido as variáveis aleatórias X_i convirjam para X.
- Situações como essa podem surgir quando consideramos *n* observações independentes de uma variável aleatória X, por exemplo, X_1, ..., X_n. Poderemos estar interessados na média aritmética dessas observações

$$\bar{X}_n = (X_1 + \ldots + X_n)/n,$$

que é também uma variável aleatória. Seja $\bar{F}_n$ a fd de $\bar{X}_n$. Poderá interessar saber o que ocorre à distribuição de $\bar{X}_n$ quando *n* se torna grande. O teorema seguinte auxilia na resolução desse problema.

10.6 Sequencias de Variáveis Aleatórias

- **Teorema:**

  Seja $X_1, \ldots, X_n, \ldots$, uma sequência de variáveis aleatórias com fd $F_1, \ldots, F_n, \ldots$, e fgm $M_1, \ldots, M_n, \ldots$. Suponha-se que

  $$\lim_{n\to\infty} M_n(t) = M(t),$$

  em que $M(0) = 1$. Então, $M(t)$ é a fgm da variável aleatória X, cuja fd F é dada pelo $\lim_{n\to\infty} F_n(t)$.

- **Comentários:**

  O teorema diz que para obter a distribuição limite F procurada é ***suficiente*** estudar as funções geratrizes de momentos das variáveis aleatórias sob exame. Obteremos a forma limite da sequência $M_1, \ldots, M_n, \ldots$, digamos $M(t)$. Em virtude da unicidade das fgm, existirá apenas uma distribuição de probabilidade correspondente à fgm $M(t)$. Assim, poderemos reconhecer M como a fgm de uma distribuição conhecida (tal como a normal, a de Poisson etc.).