PROBABILIDADE E PROCESSOS ESTOCÁSTICOS

DINTER UEA-UFPE

VII LISTA

1) Estratégias de amostragem.

$M_{A_n} := \frac{1}{n}\sum_{i=1}^{n} X_i$, com X_i não correlacionados par-a-par.

$M_{B_n} := \frac{1}{2n}\sum_{i=1}^{n} X_i$, com amostras adjacentes correlacionadas,

E{X_1X_2}= E{X_2X_3}= E{X_3X_4}=...= E{X_iX_{i+1}}=ρ e E{X_iX_j}=0 se $i \neq j \pm 1$.

Assuma X_i de média nula e variância unitária.

   a) Determine a variância var(M_{An})
   b) Determine a variância var(M_{Bn})
   c) Se ρ=1/2, encontre $\Delta\sigma_n^2 := \sigma_{M_{An}}^2 - \sigma_{B_{An}}^2$. Dê uma interpretação sobre o que os resultados implicam na eficiência das duas técnicas.

2) Seja {Xi} uma seqüência de v.a.'s i.i.d. com $E\{X_i^4\} < \infty$. Esta seqüência pode ser interpretada como uma seqüência de realizações de uma variável aleatória X obtida por repetidas execuções de um mesmo experimento. Na prática, frequentemente tentamos estimar a média E{X}=m usando a média amostral $M_n := \frac{1}{n}\sum_{i=1}^{n} X_i$. Observamos que $E\{M_n\} = m$ de forma que a estimativa é não enviezada. Ademais, com base na lei fraca dos grandes números, sabe-se que $M_n \xrightarrow{P} m$. Neste problema, queremos investigar um estimador da variância $E\{(X-m)^2\} = \sigma^2$. Defina a variância amostral por:

$V_n := \frac{1}{n}\sum_{i=1}^{n} (X_i - M_n)^2$.

   a) V_n é um estimador não-enviezado de σ^2? É assintoticamente não enviezado?
   b) Prove que $V_n \xrightarrow{P} \sigma^2$.

3) Mostre que $E\left(\frac{n_A}{n}\right) = p$ e que $\mathrm{var}\left(\frac{n_A}{n}\right) = \frac{p(1-p)}{n}$, i.e. o valor esperado da freqüência relativa de ocorrência de um dado evento é igual a probabilidade de ocorrência de tal evento; e a variância da

freqüência relativa varia inversamente com o número de amostras aleatórias independentes.

4) a) Na questão anterior, mostre que a variância atinge o seu valor máximo quando p=1/2, i.e., mostre que $\mathrm{var}\left(\frac{n_A}{n}\right) \leq \frac{1}{4n}$. b) Mostre que se segue, portanto, que: $\Pr\left(\left|\frac{n_A}{n} - p\right| \geq \varepsilon\right) \leq \frac{1}{4n\varepsilon^2}$.

5) Se $\hat{m} := \sum_{i=1}^{n} a_i X_i$ é um estimador linear não-polarizado da média, calculado a partir das amostras, mostre que os coeficientes a_i devem satisfazer $\sum_{i=1}^{n} a_i = 1$.

6) Considere o estimador linear $\hat{m}$ definido anteriormente. Mostre que se os coeficientes a_i são ajustados de maneira que a variância de $\hat{m}$ seja minimizada quando as amostras aleatórias são não-correlacionadas par a par, então $a_i = \frac{1}{n}$, i=1,2.,...,n.

7) Sejam X e Y um par de variáveis aleatórias com função densidade de probabilidade conjunta $f_{X,Y}$ e uma função g(X) mensurável e de segundo momento finito; Mostre que se usamos g(X) como um preditor de Y, $\hat{Y} = g(X)$, então $E\left\{\left(Y - \hat{Y}\right)^2\right\} = E\left[Y^2\right] - 2E\left[g_0(X)g(X)\right] + E\left[g^2(X)\right]$, em que $g_0(X) := E(Y \mid X = x) = \int_{-\infty}^{+\infty} y f_{Y \mid X}(y \mid x)dy$.

8) Tomemos X e Y da questão anterior, com $g(X)$ e $g_0(X)$ definidos como anteriormente. a) mostre que se tomamos $g_0(X)$ como o preditor de Y, i.e., $\hat{Y}_0 := g_0(X) = E(Y \mid X = x)$, então $E\left[\left(Y - \hat{Y}_0\right)^2\right] = E[Y^2] - E\left[g_0^2(X)\right]$. b) Mostre que $E\left[\left(Y - \hat{Y}\right)^2\right] \geq E\left[\left(Y - \hat{Y}_0\right)^2\right]$ para qualquer função mensurável g que tenha segundo momento finito.