Lecture 40 - 12/05

Monday, December 05, 2022 10:16 AM

Review.

① Minimum mean square error estimators of random variable Y, after observing random variable X

. Constant estimator: $\hat{s}^* = E[Y]$, error $= \sigma_Y^2$

. Linear estimator: $L^*(X) = \hat{E}[Y|X] = \mu_Y + \sigma_Y \rho_{X,Y}\left(\frac{X-\mu_X}{\sigma_X}\right)$, error $= \sigma_Y^2(1-\rho_{X,Y}^2)$

. Unconstrained estimator: $g^*(X) = E[Y|X]$, error $= \sigma_Y^2 - E[(E[Y|X])^2]$

② Law of large numbers

Suppose that $X_1, X_2, \dots, X_n$ are uncorrelated, each one has the same mean μ, and their variance is bounded by C. Let $S_n = X_1 + X_2 + \dots + X_n$. Then, for any $\delta > 0$, we have

$$P\left(\left\{\left|\frac{S_n}{n} - \mu\right| \geq \delta\right\}\right) = P\left(\left\{w \in \Omega : \left|\frac{S_n(w)}{n} - \mu\right| \geq \delta\right\}\right) \leq \frac{C}{n\delta^2} \xrightarrow{\text{as } n\to\infty} 0$$

③ Central limit theorem

Suppose that $X_1, X_2, \dots, X_n$ are independent, identically distributed, each with mean μ and variance σ^2. Let $S_n = X_1 + X_2 + \dots + X_n$. Then, for any c, we have

$$\lim_{n\to\infty} P\left(\left\{\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq c\right\}\right) = \lim_{n\to\infty} P\left(\left\{\tilde{S}_n \leq c\right\}\right) = \Phi(c)$$

where $\tilde{S}_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} = \frac{S_n - E[S_n]}{\sqrt{Var(S_n)}}$ is the standardized version of S_n.

④ Jointly Guassian random variables

We say random variables X and Y are jointly guassian random variables if any linear combination of X and Y are guassian random variables, i.e., for any $a, b \in \mathbb{R}$, $aX + bY$ is a guassian random variables.

. class of bivariate normal pdfs are preserved via linear transformation

. setting $a=0$ and $b=1$, implies that $aX + bY = Y$ is $\mathcal{N}(\mu_Y, \sigma_Y^2)$. Similarly, X is $\mathcal{N}(\mu_X, \sigma_X^2)$.

. If $\rho = 0$ then $f_{X,Y}(u,v) = f_X(u) f_Y(v)$. Hence uncorrelated jointly guassian random variables are independent.

. Let $Z = aX + bY$. Then X and Z are jointly guassian.

. If X and Y are guassian, it does not mean that they are jointly guassian.

Proposition: Suppose that X & Y have bivariate normal pdf with parameters $\mu_X, \mu_Y, \sigma_X, \sigma_Y$ and ρ.

$$f_{X,Y}(u,v) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(\frac{\left(\frac{u-\mu_X}{\sigma_X}\right)^2 + \left(\frac{v-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\left(\frac{u-\mu_X}{\sigma_X}\right)\left(\frac{v-\mu_Y}{\sigma_Y}\right)}{2(1-\rho^2)}\right)$$

(a) X has $\mathcal{N}(\mu_X, \sigma_X^2)$ distribution, Y has $\mathcal{N}(\mu_Y, \sigma_Y^2)$ distribution.

(b) for any $a, b \in \mathbb{R}$, $aX + bY$ is a guassian random variable, i.e., X and Y are jointly guassian.

(a) X has $\mathcal{N}(\mu_X, \sigma_X^2)$ distribution, Y has $\mathcal{N}(\mu_Y, \sigma_Y^2)$ distribution.

(b) For any $a, b \in \mathbb{R}$, $aX+bY$ is a guassion random variable, i.e., X and Y are jointly guassion.

(c) $\rho_{X,Y} = \rho$

(d) if $\rho = 0$, then they are independent

(e) For estimation of Y from X, $L^*(X) = g^*(X)$. Equivalently, $E[Y|X] = \hat{E}[Y|X]$, i.e., the best linear estimator of Y using X, is the best unconstrained estimator.

(d) The conditional distribution of Y given $X=u$ is $N(\hat{E}[Y|X=u], \sigma_e^2)$ where σ_e^2 is MSE for $\hat{E}[Y|X]$.

proof: We will prove it for $\mu_X = \mu_Y = 0$ and $\sigma_X^2 = \sigma_Y^2 = 1$ as the general case follows by a linear transformation.

$$f_{X,Y}(u,v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{u^2+v^2-2\rho uv}{2(1-\rho^2)}\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \cdot \left(\frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(v-\rho u)^2}{2(1-\rho^2)}\right)\right)$$

For fixed u, this is pdf of guassion random variable with distribution $\mathcal{N}(\rho u, 1-\rho^2)$

(a) $$f_X(u) = \int_{-\infty}^{+\infty} f_{X,Y}(u,v)\, dv = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \cdot \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(v-\rho u)^2}{2(1-\rho^2)}\right) dv = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

similarly $f_Y(v) = \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}$

(b) Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for c, d s.t. $ad - bc \neq 0$. Define $\begin{pmatrix} W \\ Z \end{pmatrix} = A \begin{pmatrix} X \\ Y \end{pmatrix}$. Recall that class of bivariate normal pdfs are preserved via linear transformation. Hence $\begin{pmatrix} W \\ Z \end{pmatrix}$ has bivariate normal pdf and by part (a), $W = aX + bY$ has normal distribution.

(c) & (d) Notice that $f_{X,Y}(u,v) = f_X(u) f_Y(v|u)$. Hence, by part (a)

$$f_{Y|X}(v|u) = \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(v-\rho u)^2}{2(1-\rho^2)}\right)$$

In particular, $E[Y|X=u] = \rho u$. Hence, $g^*(X) = \rho X$. Moreover,

$$\begin{aligned}
\rho_{X,Y} = E[XY] &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} uv\, f_{X,Y}(u,v)\, dv\, du \\
&= \int_{-\infty}^{+\infty} u f_X(u) \int_{-\infty}^{+\infty} v f_{Y|X}(v|u)\, dv\, du \\
&= \int_{-\infty}^{+\infty} u f_X(u) \cdot \rho u\, du = \rho
\end{aligned}$$

4.37. **[Transforming joint Gaussians to independent random variables]**
Suppose X and Y are jointly Gaussian such that X is $N(0,9)$, Y is $N(0,4)$, and the correlation coefficient is denoted by ρ. The solutions to the questions below may depend on ρ and may fail to exist for some values of ρ.

(a) For what value(s) of a is X independent of $X + aY$?
(b) For what value(s) of b is $X + Y$ independent of $X - bY$?
(c) For what value(s) of c is $X + cY$ independent of $X - cY$?
(d) For what value(s) of d is $X + dY$ independent of $(X - dY)^3$?

4.30. **[Law of Large Numbers and Central Limit Theorem]**
A fair die is rolled n times. Let $S_n = X_1 + X_2 + \ldots + X_n$, where X_i is the number showing on the i^{th} roll. Determine a condition on n so the probability the sample average $\frac{S_n}{n}$ is within 1% of the mean μ_X, is greater than 0.95. (Note: This problem is related to Example 4.10.6.)

(a) Solve the problem using the form of the law of large numbers based on the Chebychev inequality (i.e. Proposition 4.10.1 in the notes).
(b) Solve the problem using the Gaussian approximation for S_n, which is suggested by the CLT. (Do not use the continuity correction, because, unless $3.5n \pm (0.01)n\mu_X$ are integers, inserting the term 0.5 is not applicable).