Class Project 4, Covariance and Correlation, Limit Theorems

ECE 313

Probability with Engineering Applications

Lecture 18

Professor Ravi K. Iyer

Dept. of Electrical and Computer Engineering

University of Illinois at Urbana Champaign

Today's Topics

- Class Project 4:
- Hypothesis Testing Example:
  - Continuous type observations
- Covariance and Correlation
- Limit Theorems

Class Project 4

- Two biomedical signals, the blood pressure (ABP) and the heart rate (HR), are measured to detect the abnormalities of a patient in an intensive care unit (ICU).
- Assume that the blood pressure sensor outputs a value *X* and the heart rate sensor outputs a value *Y*.
- *X* and *Y* outputs have possible values of {0, 1, 2}, representing different ranges of ABP and HR values, with larger numbers tending to indicate that an abnormality is present.
- Let H_0 be the hypothesis there is no abnormality, and H_1 be the hypothesis an abnormality is present.
- The likelihood matrices for *X* and for *Y* are shown:

	$X=0$	$X=1$	$X=2$
H_1	0.1	0.3	0.6
H_0	0.8	0.1	0.1

	$Y=0$	$Y=1$	$Y=2$
H_1	0.1	0.1	0.8
H_0	0.7	0.2	0.1

Class Project 4 (Cont'd)

- Suppose, given one of the hypotheses is true, the sensors provide conditionally independent readings, so that:

$$P(X = i, Y = j \mid H_k) = P(X = i \mid H_k).P(Y = j \mid H_k) \quad for\ i, j \in \{0,1,2\}\ and\ k \in \{0,1\}$$

a) Find the likelihood matrix for the observation (*X*, *Y*), and indicate the ML decision rule. To be definite, break ties in favor of H_1.

b) Find $p_{false-alarm}$ and p_{miss} for the ML rule found in part (a).

c) Suppose, based on past experience, prior probabilities are assigned as: $(p_0, p_1) = (0.8, 0.2)$ Compute the joint probability matrix and indicate the MAP decision rule.

d) For the MAP decision rule, compute $p_{false-alarm}$,p_{miss}, and the unconditional probability of error $p_e = \pi_0 p_{false-alarm} + \pi_1 p_{miss}$.

e) Using the same priors as in part (c), compute the unconditional error probability for the ML rule from part (a). Is it smaller or larger than the p_e found for the MAP rule in part (d)?

Hypothesis Testing Example

- Suppose under hypothesis H_i; the observation X is notmally distributed with distribution $N(m_i;\sigma^2)$ distribution, for $i = 0$ or $i = 1$; where the parameters are known and satisfy: $\sigma^2 > 0$ and $m_0 < m_1$.

- Identify the ML and MAP decision rules and their associated error probabilities, $p_{false-alarm}$ and p_{miss}.
- Assume prior probabilities π_0 *and* π_1 are given where needed.

- The pdfs are given by:

$$f_i(u) = \frac{1}{\sqrt{2\pi\sigma}} \exp\{-\frac{(u-m_i)^2}{2\sigma^2}\}$$

Hypothesis Testing Example (Cont'd)

- So:

$$\begin{aligned}\Lambda(u) &= \frac{f_1(u)}{f_0(u)} \\ &= \exp\left\{-\frac{(u-m_1)^2}{2\sigma^2} + \frac{(u-m_0)^2}{2\sigma^2}\right\} \\ &= \exp\left\{\left(u - \frac{m_0+m_1}{2}\right)\left(\frac{m_1-m_0}{\sigma^2}\right)\right\}.\end{aligned}$$

- Observe that $\Lambda(X) > 1$ if only and if $X > \frac{m_0+m_1}{2}$, so the ML rule is:

$$X \begin{cases} > \gamma_{ML} & \text{declare } H_1 \text{ is true} \\ < \gamma_{ML} & \text{declare } H_0 \text{ is true.}\end{cases}$$

where $\gamma_{ML} = \frac{m_0+m_1}{2}$.

(Note that γ is used as the threshold for *X* directly, whereas τ is used for the threshold applied to the likelihood ratio).

Hypothesis Testing Example (Cont'd)

- The LRT for a general threshold τ and a general binary hypothesis testing problem with continuous-type observations is equivalent to (by taking natural logarithm of both sides of LRT):

$$\ln \Lambda(X) \begin{cases} > \ln \tau & \text{declare } H_1 \text{ is true} \\ < \ln \tau & \text{declare } H_0 \text{ is true,} \end{cases}$$

which here can be expressed as a threshold test for *X:*

$$X \begin{cases} > \left(\frac{\sigma^2}{m_1 - m_0}\right) \ln \tau + \frac{m_0 + m_1}{2} & \text{declare } H_1 \text{ is true} \\ < \left(\frac{\sigma^2}{m_1 - m_0}\right) \ln \tau + \frac{m_0 + m_1}{2} & \text{declare } H_0 \text{ is true.} \end{cases}$$

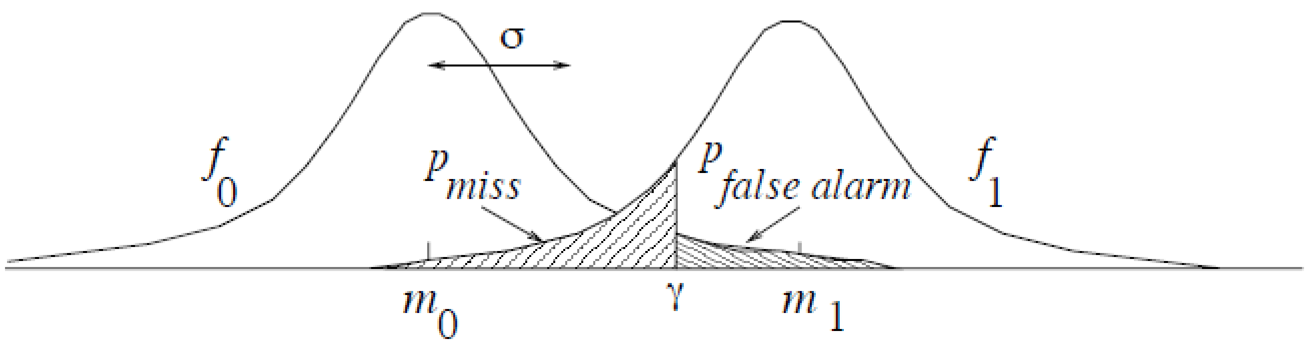
- In particular the MAP rule is obtained by letting $\tau = \dfrac{\pi_0}{\pi_1}$, so:

$$X \begin{cases} > \gamma_{MAP} & \text{declare } H_1 \text{ is true} \\ < \gamma_{MAP} & \text{declare } H_0 \text{ is true.} \end{cases}$$

where $\gamma_{MAP} = \left(\dfrac{\sigma^2}{m_1 - m_0}\right) \ln\left(\dfrac{\pi_0}{\pi_1}\right) + \dfrac{m_0 + m_1}{2}$

Hypothesis Testing Example (Cont'd)

- Therefore, we shall examine the error probabilities for a test of that form. The error probabilities are given by the areas of the shaded regions shown below:



$$\begin{aligned} p_{\text{miss}} &= P(X < \gamma | H_1) \\ &= P\left(\frac{X - m_1}{\sigma} < \frac{\gamma - m_1}{\sigma} \middle| H_1 \right) \\ &= Q\left(\frac{m_1 - \gamma}{\sigma} \right). \end{aligned}$$

$$\begin{aligned} p_{\text{false alarm}} &= P(X > \gamma | H_0) \\ &= P\left(\frac{X - m_0}{\sigma} > \frac{\gamma - m_0}{\sigma} \middle| H_0 \right) \\ &= Q\left(\frac{\gamma - m_0}{\sigma} \right). \end{aligned}$$

$$p_e = \pi_0 p_{\text{false alarm}} + \pi_1 p_{\text{miss}}.$$

Hypothesis Testing Example (Cont'd)

- Substituting in $\gamma_{ML} = \frac{m_0 + m_1}{2}$ in the error expressions yields that error probabilities for the ML rule for this example satisfy:

$$p_{\text{false alarm}} = p_{\text{miss}} = p_e = Q\left(\frac{m_1 - m_0}{2\sigma}\right).$$

- Note that $\frac{m_0 - m_1}{\sigma}$ can be interpreted as a signal-to-noise ratio.
- The difference in the means $m_0 - m_1$ can be thought of as the difference between the hypotheses due to the signal, and is the standard deviation of the noise.

- The error probabilities for the MAP rule can be obtained by substituting in $\gamma = \gamma_{MAP}$ in the error expressions.

Covariance and Variance

- Recall $$E[X^n] = \begin{cases} \sum_{x:p(x)>0} x^n p(x), & \text{if X is discrete} \\ \int_{-\infty}^{\infty} x^n f(x)dx, & \text{if X is continuous} \end{cases}$$
- And Var(X), which is defined by $Var(X) = E[(X - E[X])^2]$
- The variance of X measures the expected square of the deviation of X from its expected value: $Var(X) = E[X^2] - (E[X])^2$

- **Covariance and Variance of Sums of Random Variables**
- The covariance of any two random variables, X and Y, denoted by $Cov(X,Y)$, is defined by

$$\begin{aligned} Cov(X,Y) &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY - YE[X] - XE[Y] + E[X]E[Y]] \\ &= E[XY] - E[Y]E[X] - E[X]E[Y] + E[X]E[Y] \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

- If X and Y are independent then it follows that $Cov(X,Y)=0$

Covariance and Variance Example

- In general it can be shown that a positive value of *Cov(X,Y)* is an indication that *Y* tends to increase as *X* does, whereas a negative value indicates that *Y* tends to decrease as *X* increases.

- **Example:** The joint density function of *X,Y* is:

$$f(x,y) = \frac{1}{y} e^{-(y+x/y)}, \; 0 < x, y < \infty$$

a) Verify that the preceding is a joint density function.

b) Find *Cov(X,Y)*.

- To show that *f(x,y)* is a joint density function we need to show it is nonnegative, which is immediate and that $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dydx = 1$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dy\,dx = \int_{0}^{\infty}\int_{0}^{\infty} \frac{1}{y} e^{-(y+x/y)} dy\,dx$$

$$= \int_{0}^{\infty} e^{-y} \int_{0}^{\infty} \frac{1}{y} e^{-x/y} dx\,dy$$

$$= \int_{0}^{\infty} e^{-y} dy = 1$$

Covariance and Variance Example

- To obtain *Cov(X,Y)*, note that the density function of *Y* is

$$f_Y(y) = e^{-y}\int_0^\infty \frac{1}{y}e^{-x/y}dx = e^{-y}$$

- Thus *Y* is an exponential random variable with parameter 1

$E[Y]=1$

- Compute *E[X]* and *E[XY]* as follows:

$$E[X] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dy\,dx$$

$$= \int_0^\infty e^{-y}\int_0^\infty \frac{x}{y}e^{-x/y}dx\,dy$$

- Now, $\int_0^\infty \frac{x}{y}e^{-x/y}dx$ is the expected value of an exponential random variable with parameter 1/*y*, and thus is equal to *y*. Consequently,

$$E[X] = \int_0^\infty ye^{-y}dy = 1$$

Covariance and Variance

- Also $E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dy\,dx$

$$= \int_0^\infty ye^{-y}\int_0^\infty \frac{x}{y}e^{-x/y}dx\,dy$$

$$= \int_0^\infty y^2e^{-y}dy$$

- Integration by parts $(dv = e^{-y}dy, u = y^2)$ gives

$$E[XY] = \int_0^\infty y^2e^{-y}\,dy = -y^2e^{-y}\Big|_0^\infty + \int_0^\infty 2ye^{-y}dy = 2E[Y] = 2$$

- Consequently, $Cov(X,Y) = E[XY] - E[X]E[Y] = 1$

Properties of Covariance

- For any random variable X, Y, Z, and constant c, we have:
  1. Cov (X,X) = Var(X),
  2. Cov (X,Y) = Cov(Y,X),
  3. Cov (cX,Y) = cCov(X,y),
  4. Cof (X,Y+Z) = Cov(X,Y) = Cov(X,Z).

  Whereas the first three properties are immediate, the final one is easily proven as follows:

$$\begin{aligned} Cov(X, Y+Z) &= E[X(Y+Z)] - E[X]E[Y+Z] \\ &= E[XY] - E[X]E[Y] + E[XZ] - E[X]E[Z] \\ &= Cov(X,Y) + Cov(X,Z) \end{aligned}$$

- The last property generalizes to give the following result:

$$Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov(X_i, Y_j)$$

Properties of Covariance

- A useful expression for the variance of the sum of random variables can be obtained from the preceding equation

$$Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i=1}^{n}\sum_{j<i} Cov(X_i, Y_j)$$

- If $X_i, i = 1, \ldots, n$ are independent random variables, then the above equation reduces to

$$Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

Limit Theorems

- Proposition (Markov's Inequality) If X is a random variable that takes only nonnegative values, then for any value $a>0$

$$P\{X \geq a\} \leq \frac{E[X]}{a}$$

- **Proof:** X is continuous with density f.

$$\begin{aligned}
E[X] &= \int_0^\infty xf(x)dx \\
&= \int_0^a xf(x)dx + \int_a^\infty xf(x)dx \\
&\geq \int_a^\infty xf(x)dx \\
&\geq \int_a^\infty af(x)dx \\
&= a\int_a^\infty f(x)dx \\
&= aP\{X \geq a\}
\end{aligned}$$

Limit Theorems Cont'd

- A corollary, we obtain the following
- Proposition (Chebyshev's Inequality) If X is a random variable with mean μ and variance σ^2 then for any value $k>0$,

$$P\{|X-\mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

- **Proof:** Since $(X-\mu)^2$ is a nonnegative random variable, we can apply Markov's inequality (with $a=k^2$) to obtain

$$P\{(X-\mu)^2 \geq k^2\} \leq \frac{E[(X-\mu)^2]}{k^2}$$

- Since $(X-\mu)^2 \geq k^2$ if and only if $|X-\mu| \geq k$ is equivalent to

$$P\{|X-\mu| \geq k \} \leq \frac{E[(X-\mu)^2]}{k^2} = \frac{\sigma^2}{k^2}$$

- And the proof is complete

Limit Theorems Example

- The importance of Markov's and Chebyshev's inequalities is that they enable us to derive bounds on probabilities when only the mean, or both the mean and the variance, of the probability distribution are known. If the actual distribution were known, then the desired probabilities could be exactly computed, and we would not need to resort to bounds.

- Example: Suppose we know that the number of items produced in a factory during a week is a random variable with mean 500.
  a) What can be said about the probability that this week's production will be at least 1000?
  b) If the variance of a week's production is known to equal 100, then what can be said about the probability that this week's production will be between 400 and 600?

Limit Theorems Example (Cont'd)

- Let X be the number of items that will be produced in a week.
  a) By Markov's inequality,

$$P\{X \geq 1000\} \leq \frac{E[X]}{1000} = \frac{500}{1000} = \frac{1}{2}$$

  b) By Chebyshev's inequality,

$$P\{|X - 500| \geq 100\} \leq \frac{\sigma^2}{(100)^2} = \frac{1}{100}$$

  Hence,

$$P\{|X - 500| < 100\} \leq 1 - \frac{1}{100} = \frac{99}{100}$$

And so the probability that this week's production will be between 400 and 600 is at least 0.99.