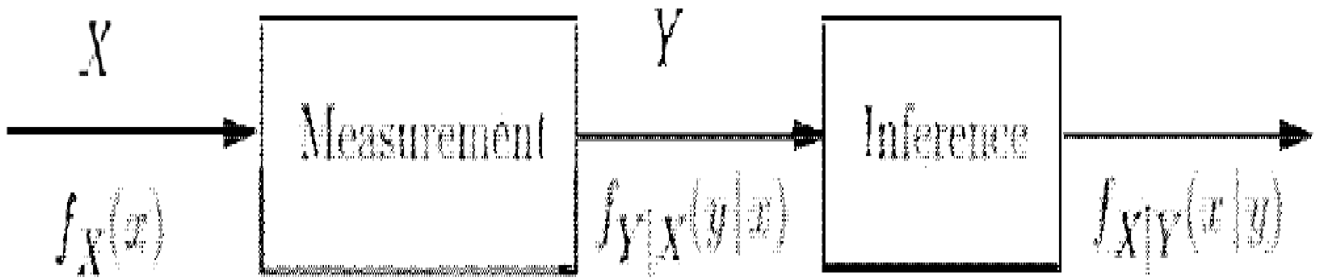
MA 6.101
Probability and Statistics

Tejas Bodas

Assistant Professor, IIIT Hyderabad

Inference problem



- X is an unobservable random variable with a known distribution.
- We only observe measurements Y that takes values according to $f_{Y|X}(y|x)$.
- Objective is to draw inference about X having seen a realization of Y i.e., Obtain $f_{X|Y}(x|y)$ using only $f_X(x)$ and $f_{Y|X}(y|x)$, both of which are known.

Bayes Rule revisited

$$P(B/A) = \frac{P(A/B)P(B)}{P(A)}$$

For continuous random variables X and Y

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)} = \frac{f_{Y|X}(y|x)f_X(x)}{\int_{-\infty}^{\infty} f_{Y|X}(y|t)f_X(t)dt}$$

For discrete random variables X and Y

$$p_{X|Y}(x|y) = \frac{p_{Y|X}(y|x)p_X(x)}{p_Y(y)} = \frac{p_{Y|X}(y|x)p_X(x)}{\sum_i p_{Y|X}(y|i)p_X(i)}$$

Example 3.19(Bertsekas)

Lifetime of a Phillips bulb is assumed to be an exponential random variable Y with parameter Λ. Λ itself is a uniform random variable over $[1, 1.5]$. You test a bulb and see that it has a lifetime of y units. What can you say about randomness of Λ having observed $Y = y$.?

- What is $f_\Lambda(\lambda)$?
- What is $f_{Y|\Lambda}(y|\lambda)$?
- What is $f_Y(y)$?
- $f_{\Lambda|Y}(\lambda|y) = \frac{2\lambda e^{-\lambda y}}{\int_1^{1.5} 2te^{-ty} dt}$ for $\lambda \in [1, 1.5]$.

Bayes Rule revisited

For discrete N and continuous random variable Y

$$P(N = n|Y = y) = \frac{f_{Y|N}(y|n)p_N(n)}{f_Y(y)} = \frac{f_{Y|N}(y|n)p_N(n)}{\sum_i f_{Y|N}(y|i)p_N(i)}$$

Equivalently

$$f_{Y|N}(y|n) = \frac{P(N = n|Y = y)f_Y(y)}{p_N(n)} = \frac{P(N = n|Y = y)f_Y(y)}{\int_{-\infty}^{\infty} P(N = n|Y = t)f_Y(t)dt}$$

Example 3.20 (Bertsekas)

- Suppose $X = 1$ w.p. p and $X = -1$ w.p. $1 - p$. While transmitting this signal, it is corrupted by a Gaussian noise $N \sim \mathcal{N}(0, 1)$. We observe $Y = X + N$. Suppose you observe $Y = y$, then show that

$$P(X = 1 | Y = y) = \frac{pe^y}{pe^y + (1-p)e^{-y}}$$

- Intuitively, this probability goes to zero as y decreases to $-\infty$ and increases to 1 as y increases to ∞.
- $P(X = 1 | Y = y) = \frac{f_{Y|X}(y|1)p_X(1)}{f_Y(y)}$
- Here $f_Y(y) = f_{Y|X}(y|1)p_X(1) + f_{Y|X}(y| - 1)p_X(1)$.
- Substitute values to obtain answer.

Law of Iterated Expectation revisited

- Recall $E[X] = E[E[X|Y]]$. What are the two expectations w.r.t ?
- Let $g(Y) = E[X|Y]$. Then

$$g(y_1) = E[X|Y = y_1] = \int_x x f_{X|Y}(x|y_1) dx$$

. So the inner expectation is w.r.t X.
- $E[X] = E[g(Y)] = \int_y g(y) f_Y(y) dy$. So the outer expectation is w.r.t Y.

$$\boxed{E[X] = E_Y\left[E_X[X|Y]\right]}$$

Law of Iterated Expectation revisited

$$E[X] = E_Y\left[E_X[X|Y]\right]$$

- What is $E[Xg(Y)|Y]$?
- Note that $E[Xg(Y)|Y = y_1] = g(y_1)E[X|Y = y_1]$.
- Therefore $E[Xg(Y)|Y] = g(Y)E[X|Y]$. In general, we have the following pull through property

$$E[h(X)g(Y)|Y] = g(Y)E[h(X)|Y].$$

Law of Iterated Expectation revisited

$$E[X] = E_Y\left[E_X[X|Y]\right]$$

- If X and Y are independent, what is $E[X|Y]$?
- Since X and Y are independent, $E[X|Y = y] = E[X]$ for all y.
- This means $g(Y) = E[X|Y]$ always takes the value of $E[X]$.

In fact, when X and Y are independent, we have

$$E[g(X)|Y] = E[g(X)].$$

Variance of sum of random variables

- Let $X_1, X_2, \ldots X_n$ be possibly dependent and non-identical random variables.
- Lets say you know the joint pdf/pmf for every pair of random variables from this collection.
- AIM: Calculate $Var(Z)$ where $Z = \sum_{i=1}^{n} a_i X_i$ for some scalars a_i.

Variance of sum of random variables

- Recall $Var(X) = E[X - E[X]]^2 = E[X^2] - E[X]^2$.
- Also recall $Cov(X, Y) = E[XY] - E[X]E[Y]$.
- Following properties of covariance follow (HW)

1. $Cov(X, X) = Var(X)$
2. If X, Y are independent, $Cov(X, Y) = 0$.
3. $Cov(X, Y) = Cov(Y, X)$
4. $Cov(aX, Y) = aCov(X, Y)$
5. $Cov(X + a, Y) = Cov(X, Y)$
6. $Cov(X + Z, Y) = Cov(X, Y) + Cov(Z, Y)$
7. $Cov\left(\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j Cov(X_i, Y_j)$

Variance of sum of random variables

- AIM: Calculate $Var(Z)$ where $Z = \sum_{i=1}^{n} a_i X_i$ for some scalars a_i.
- $Var(Z) = Cov(Z, Z)$ and therefore

$$
\begin{aligned}
Cov\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{n} a_j X_j\right) &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_i a_j Cov(X_i, X_j) \\
&= \sum_{i=1}^{n} a_i^2 Var(X_i) \\
&+ \sum_{(i,j):i\neq j} a_i a_j Cov(X_i, X_j)
\end{aligned}
$$

$$
\boxed{Var(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 Var(X_i) + \sum_{(i,j):i\neq j} a_i a_j Cov(X_i, X_j)}
$$

Variance of sum of random variables

$$Var(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 Var(X_i) + \sum_{(i,j):i\neq j} a_i a_j Cov(X_i, X_j)$$

- Show that $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$
- Now if $X_i's$ are independent, what is $Var(Z)$?
- Let $\{X_i, i = 1, 2, \ldots n\}$ be i.i.d and consider $S_n = \frac{\sum_{i=1}^{n} X_i}{n}$.
- Show that $Var(S_n) = \frac{Var(X)}{n}$