CMPE 320: Probability, Statistics, and Random Processes

Lecture 19: Conditional expectation, Independence, and Bayes rule

Spring 2018

Seung-Jun Kim

Announcement

- Special review sessions
  - For those who feel extra shaky about basic problem solving
  - Will mainly review midterm 1 and 2 problems
  - 2~3 times before the final, once a week, taught by the TA

- If interested, please e-mail by tomorrow (5/1) to the TA (lee43@umbc.edu)
  - Will try to find the common time slot online

Conditional expectation

- Instead of ordinary PDFs, use conditional PDFs
- Analogous formulas to discrete RVs, except sums are replaced by integrals

$$E[X \mid A] = \int_{-\infty}^{\infty} x f_{X|A}(x)\, dx$$

$$E[X \mid Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\, dx$$

For a function $g(X)$

$$E[g(X) \mid A] = \int_{-\infty}^{\infty} g(x) f_{X|A}(x)\, dx$$

$$E[g(X) \mid Y=y] = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y)\, dx$$

Total expectation

- Let $A_1, A_2, \dots, A_n$ be disjoint events that form a partition of Ω with $P(A_i) > 0$

Recall the total probability theorem for PDF

$$f_X(x) = \sum_{i=1}^{n} P(A_i) f_{X|A_i}(x)$$

Multiply both sides by x and integrate

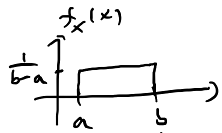
$$\int_{-\infty}^{\infty} x f_X(x)\, dx = \int_{-\infty}^{\infty} x \sum_{i=1}^{n} P(A_i) f_{X|A_i}(x)\, dx = \sum_{i=1}^{n} P(A_i) \int_{-\infty}^{\infty} x f_{X|A_i}(x)\, dx$$

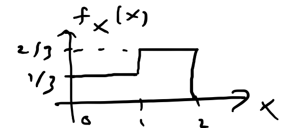
$$\boxed{E[X] = \sum_{i=1}^{n} P(A_i) E[X \mid A_i]}$$

Similarly, $\boxed{E[X] = \int_{-\infty}^{\infty} f_Y(y) E[X \mid Y=y]\, dy}$ ← start from $f_X(x) = f_{X|Y}(x|y) f_Y(y)$

Example 3.17. Mean and Variance of a Piecewise Constant PDF. Suppose that the random variable X has the piecewise constant PDF

$$f_X(x) = \begin{cases} 1/3, & \text{if } 0 \le x \le 1, \\ 2/3, & \text{if } 1 < x \le 2, \\ 0, & \text{otherwise,} \end{cases}$$





$E(X) = \frac{a+b}{2}$

$E(X^2) = \int_a^b x^2 \frac{1}{b-a} dx = \frac{a^2+ab+b^2}{3}$

Compute E[X] and var(X) by using the total expectation theorem considering the events

$A_1 = \{X \text{ lies in the first interval } [0,1]\}$,

$A_2 = \{X \text{ lies in the second interval } (1,2]\}$.

$$E[X] = \underbrace{E[X|A_1]}_{} P(A_1) + \underbrace{E[X|A_2]}_{} P(A_2)$$

$f_{X|A_1}$ is uniform in $[0,1]$ $\qquad$ $f_{X|A_2}$ uniform in $[1,2]$

$$= \frac{1}{2}\cdot\frac{1}{3} + \frac{3}{2}\cdot\frac{2}{3} = \frac{1}{6} + 1 = \frac{7}{6}$$

$$E[X^2] = E[X^2|A_1]P(A_1) + E[X^2|A_2]P(A_2)$$

$$= \frac{0^2+0\cdot1+1^2}{3}\cdot\frac{1}{3} + \frac{1^2+1\cdot2+2^2}{3}\cdot\frac{2}{3} = \frac{5}{3}$$

$$\text{var}(X) = E(X^2) - (E[X])^2 = \frac{5}{3} - \left(\frac{7}{6}\right)^2 = \frac{11}{36}$$

Conditional expectation with multiple RVs

$$E[g(X,Y) \mid Y=y] = \int_{-\infty}^{\infty} g(x,y) f_{X|Y}(x|y)\,dx$$

$$E[g(X,Y)] = \int_{-\infty}^{\infty} E[g(X,Y) \mid Y=y] f_Y(y)\,dy$$

Independence

- Two RVs X and Y are independent if the joint PDF is the product of marginals

$$\boxed{f_{X,Y}(x,y) = f_X(x)\, f_Y(y)} \quad \text{for all } x, y$$

Compare with $f_{X,Y}(x,y) = f_{X|Y}(x|y) f_Y(y)$

$$\boxed{f_{X|Y}(x|y) = f_X(x)} \quad \text{for all } x, y \text{ with } f_Y(y) > 0$$

Similarly, $\boxed{f_{Y|X}(y|x) = f_Y(y)}$ for all x, y with $f_X(x) > 0$

- For more than 2 RVs

$$f_{X,Y,Z}(x,y,z) = f_X(x)\, f_Y(y)\, f_Z(z) \quad \text{for all } x, y, z$$

Example 3.18. Independent Normal Random Variables. Let X and Y be independent normal random variables with means μ_x, μ_y, and variances σ_x^2, σ_y^2, respectively. What is their joint PDF? Also, plot the contour $\{(x,y) | f_{X,Y}(x,y) = c\}$.

Since X and Y are independent

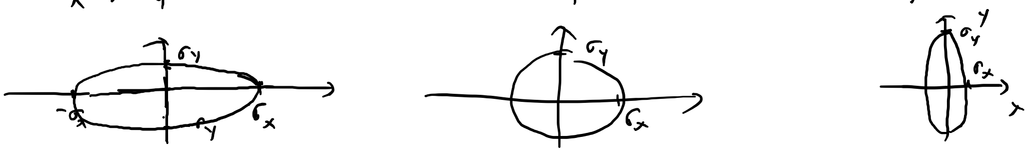
$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} \cdot \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}}$$

$$= C\, e^{-\left[\frac{(x-\mu_x)^2}{2\sigma_x^2} + \frac{(y-\mu_y)^2}{2\sigma_y^2}\right]}$$

Setting $f_{X,Y}(x,y) = c$, take log

$$\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} = c', \quad \text{consider } c' = 1,\ \mu_x = \mu_y = 0$$

If $\sigma_x^2 > \sigma_y^2$ $\qquad$ $\sigma_x^2 = \sigma_y^2$ $\qquad$ $\sigma_y^2 > \sigma_x^2$



Independence and CDF

- If X and Y are independent, so are $\{X \in A\}$ and $\{Y \in B\}$

$$P(X \in A, Y \in B) = \int_A \int_B f_{X,Y}(x,y)\,dy\,dx$$
$$= \int_A \int_B f_X(x) f_Y(y)\,dy\,dx \leftarrow X, Y: \text{independent}$$
$$= \int_A f_X(x)\,dx \int_B f_Y(y)\,dy$$
$$= P(X \in A)\, P(Y \in B)$$

Let $A = (-\infty, x]$, $B = (-\infty, y]$

$$P(X \le x, Y \le y) = P(X \le x)\, P(Y \le y)$$

$$\boxed{F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y)} \quad \text{for all } x, y$$

True even for discrete RVs, also useful for mixed cases
e.g. X: continuous, Y: discrete

Independence and expectation

- If X and Y are independent,

E[XY] = $E[X]\,E[Y]$

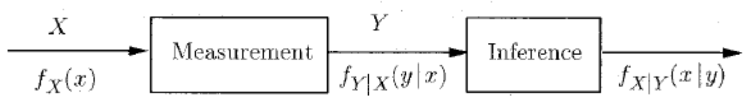
Same as discrete RVs

var(X + Y) = $var(X) + var(Y)$

- If X and Y are independent, so are g(X) and h(Y) for any functions g and h

$$E(g(X)\,h(Y)) = E(g(X))\,E(h(Y))$$

Continuous Bayes' rule



X : Some unobserved phenomenon $\sim f_X(x)$

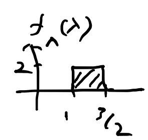
Y : Noisy measurement related to X via $f_{Y|X}(y|x)$

Want to infer about X based on $\{Y = y\}$

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

Example 3.19. A light bulb produced by the General Illumination Company is known to have an exponentially distributed lifetime Y. However, the company has been experiencing quality control problems. On any given day, the parameter λ of the PDF of Y is actually a random variable, uniformly distributed in the interval $[1, 3/2]$. We test a light bulb and record its lifetime. What can we say about the underlying parameter λ?

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y}, & y \ge 0 \\ 0, & \text{other.} \end{cases}$$



$$f_\Lambda(\lambda) = \begin{cases} 2, & \lambda \in [1, 3/2] \\ 0, & \text{otherwise} \end{cases}, \quad f_{Y|\Lambda}(y|\lambda) = \begin{cases} \lambda e^{-\lambda y}, & \text{if } y \ge 0 \\ 0, & \text{otherwise} \end{cases}$$

$$f_{\Lambda|Y}(\lambda|y) = \frac{f_{Y|\Lambda}(y|\lambda)\, f_\Lambda(\lambda)}{f_Y(y)} = \frac{f_{Y|\Lambda}(y|\lambda)\, f_\Lambda(\lambda)}{\int_{-\infty}^{\infty} f_{Y|\Lambda}(y|\lambda)\, f_\Lambda(\lambda)\, d\lambda}$$

$$= \begin{cases} \dfrac{\lambda e^{-\lambda y} \cdot 2}{\int_1^{3/2} \lambda e^{-\lambda y} \cdot 2\, d\lambda}, & 1 \le \lambda \le 3/2 \\ 0, & \text{otherwise} \end{cases}, \quad y \ge 0$$

Inference on an event based on continuous RV

- Want to infer about an event A based on observation Y = y
  - Binary signal corrupted by normally distributed noise
  - Medical diagnosis from continuous measurements (temperature, blood counts, etc.)

Given : $P(A)$, $f_{Y|A}(y)$ and $f_{Y|A^c}(y)$

Want : $P(A|Y=y) = \cancel{\dfrac{P(A, Y=y)}{P(Y=y)}}$ $\frac{0}{0}$

$$\approx \frac{P(A, y \le Y \le y+\delta)}{P(y \le Y \le y+\delta)}$$

δ : small

$P(y \le Y \le y+\delta) = \int_y^{y+\delta} f_Y(t)\,dt$

$$\approx \frac{P(y \le Y \le y+\delta | A) P(A)}{f_Y(y)\,\delta} = \frac{f_{Y|A}(y)\cdot \cancel{\delta} \cdot P(A)}{f_Y(y)\cdot \cancel{\delta}} = \frac{f_{Y|A}(y)P(A)}{f_Y(y)}$$

$$= \boxed{\frac{f_{Y|A}(y)P(A)}{f_{Y|A}(y)P(A) + f_{Y|A^c}(y)P(A^c)}} \leftarrow \text{total prob.}$$

Inference about a discrete RV

discrete RV

Apply the previous result to events of the form $\{N = n\}$

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

$f_{Y|\{N=n\}}(y)$

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

If Y is discrete

$$P(N=n | Y=k) = \frac{P(N=n, Y=k)}{P(Y=k)} = \frac{P(Y=k | N=n) P(N=n)}{P(Y=k)}$$

Example 3.20. Signal Detection. A binary signal S is transmitted, and we are given that $\mathbf{P}(S=1)=p$ and $\mathbf{P}(S=-1)=1-p$. The received signal is $Y=N+S$, where N is normal noise, with zero mean and unit variance, independent of S. What is the probability that $S=1$, as a function of the observed value y of Y?

$$P(S=1 \mid Y=y) = \frac{f_{Y|\{S=1\}}(y)\, P(S=1)}{f_Y(y)}$$

$$= \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{(y-1)^2}{2}} \cdot p}{\frac{1}{\sqrt{2\pi}} e^{-\frac{(y-1)^2}{2}} \cdot p + \frac{1}{\sqrt{2\pi}} e^{-\frac{(y+1)^2}{2}} \cdot (1-p)}$$

$f_{Y|\{S=1\}}(y) =$ Normal w. mean 1, var. 1 $= \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-1)^2}{2}}$

$f_{Y|\{S=-1\}}(y) =$ Normal w. mean -1, var. 1 $= \frac{1}{\sqrt{2\pi}} e^{-\frac{(y+1)^2}{2}}$

$$f_Y(y) = f_{Y|\{S=1\}}\, P(S=1) + f_{Y|\{S=-1\}}\, P(S=-1)$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-1)^2}{2}} \cdot p + \frac{1}{\sqrt{2\pi}} e^{-\frac{(y+1)^2}{2}} \cdot (1-p)$$

Inference based on discrete observations

$$f_{Y|A}(y) = \frac{P(A \mid Y=y)\, f_Y(y)}{P(A)}$$

$$= \frac{P(A \mid Y=y)\, f_Y(y)}{\int_{-\infty}^{\infty} P(A \mid Y=y)\, f_Y(y)\, dy}$$

When $\{N=n\}$

$$f_{Y|N}(y \mid n) = \frac{P(N=n \mid Y=y)\, f_Y(y)}{P(N=n)} = \frac{P(N=n \mid Y=y)\, f_Y(y)}{P_N(n)}$$