Recall

Conditional stuff for continuous rvs

Let X be a continuous rv; A be an event with $\mathbb{P}(\{A\}) > 0$

Have $f_{X|A}(x)$.

For X, Y both continuous rvs have joint pdf $f_{X,Y}(x,y)$

Defined conditional pdf of X given $Y=y$ via

$$f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \leftarrow \text{derived via limiting process}$$

(Generalizes to >2 rvs)

In both cases, conditional pdf is a "legit pdf" - i.e it is ≥ 0 and integrates to 1

$$\int_{-\infty}^{+\infty} f_{X|A}(x)\,dx = 1 \quad \forall A \quad , \quad \int_{-\infty}^{+\infty} f_{X|Y}(x|y)\,dx = 1 \quad \forall y$$

Conditional Expected Values

Given X, A

$$\mathbb{E}(X|A) = \int_{-\infty}^{+\infty} x\, f_{X|A}(x)\,dx$$

Given X, Y

$$\mathbb{E}(X|\{Y=y\}) = \int_{-\infty}^{+\infty} x\, f_{X|Y}(x|y)\,dx \quad \forall y$$

Expected Value Rule

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

$$\mathbb{E}(g(X) \mid Y = y) = \int_{-\infty}^{+\infty} g(x) f_{X|Y}(x|y)\, dx \quad \forall y$$

Recall the Total Probability-type resuls

- If events $A_1, A_2, \dots, A_n$ have >0 probability and partition Ω, then
  - $f_X(x) = \sum_{k=1}^{n} f_{X|A_k}(x)\, \mathbb{P}(A_k)$
  - $f_X(x) = \int_{-\infty}^{+\infty} f_{X|Y}(x|y) f_Y(y)\, dy$

From these follows

Total Expectation Theorems

$$\mathbb{E}(X) = \sum_{k=1}^{n} \mathbb{E}(X \mid A_k)\, \mathbb{P}(A_k)$$

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} \mathbb{E}(X \mid Y = y) f_Y(y)\, dy$$

What about independence?

Super General Definition: X and Y are independent when every pair of events of form $\{X \in V\}$ and $\{Y \in W\}$ are independent events.

Earlier we saw that when X,Y discrete, they are independent iff

$$p_{X,Y}(x,y) = p_X(x)p_Y(y) \quad \forall x,y$$

For **ANY** pair X,Y (continuous, discrete, or "one and one")

requires proof

$$X,Y \text{ independent} \iff F_{X,Y}(x,y) = F_X(x)F_Y(y) \quad \forall x,y$$

$\Downarrow$ implies (take $\partial/\partial x, \partial/\partial y$)

For X,Y **BOTH** continuous w/ densities $f_X(x), f_Y(y)$; $f_{X,Y}(x,y)$:

$$X,Y \text{ independent} \iff f_{X,Y}(x,y) = f_X(x)f_Y(y) \quad \forall x,y$$

Comment: When X,Y independent, we have

- $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$
- $\mathbb{E}(g(X)h(Y)) = \mathbb{E}(g(X))\mathbb{E}(h(Y))$
- $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Example - Gaussian Independence

- $X \sim \text{Gaussian}(\mu_1, \sigma_1)$
- $Y \sim \text{Gaussian}(\mu_2, \sigma_2)$

Assume independent - find $f_{X,Y}(x,y)$

$$f_{X,Y}(x,y) = f_X(x) f_Y(y)$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\right)\left(\frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}\right)$$

$$= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\left(\frac{(x-\mu_1)^2}{2\sigma_1^2} + \frac{(y-\mu_2)^2}{2\sigma_2^2}\right)}$$

Continuous Baye's Rule

Typical Setting:

$$X \longrightarrow \underset{f_{Y|X}(y|x)}{\boxed{\text{Measure}}} \longrightarrow Y \longrightarrow \underset{f_{X|Y}(x|y)}{\boxed{\text{infer}}} \longrightarrow$$ "post measurement of y" density of X

Idea: - Have a grip on $f_X(x)$

- Have a good model for $f_{Y|X}(y|x)$ $\forall\, x,y$
- Want $f_{X|Y}(x,y)$

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

Continuous Baye's Rule

Example - Λ on Some Interval

Say interval $= [1, 3/2]$

and

$$f_\Lambda(\lambda) = \begin{cases} 2 & , \lambda \in [1, 3/2] \\ 0 & , \text{else} \end{cases}$$

X is exponential w/ rate parameter Λ

Λ = quality of a lightbulb (higher quality $\Leftrightarrow$ lower Λ)

Want to infer about Λ given observations of X:

- Have $f_\Lambda(\lambda)$
- Have $f_{X|\Lambda}(x|\lambda)$ $\rightarrow$ $\begin{cases} \lambda e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$

$$f_{\Lambda|X}(\lambda|x) = \frac{f_{X|\Lambda}(x|\lambda)\, f_\Lambda(\lambda)}{\int_{-\infty}^{+\infty} f_{X|\Lambda}(x|\lambda)\, f_\Lambda(\lambda)\, d\lambda} = \begin{cases} \dfrac{2\lambda e^{-\lambda x}}{2\int_1^{3/2} \lambda e^{-\lambda x}\, d\lambda} & , \lambda \in [1, 3/2] \\ 0 & , \text{else} \end{cases}$$

If x larger $\rightarrow$ $f_{\Lambda|X}(\lambda|x)$ more focused near $\lambda = 1$

A different setting: the unobserved thing you want to infer about is discrete and the observed thing is continuous.

Example –

Say you have an event A. Have $\mathbb{P}(A)$, prior probability of A. Observe a continuous rv Y. Have a good model for $f_{Y|A}(y)$

Want: $\mathbb{P}(A \mid \{Y=y\})$. ← since $\mathbb{P}(Y=y)=0$ we must define this using a limiting process (as in Lec 19)

$$\mathbb{P}(A \mid \{Y=y\}) = \lim_{\delta \to 0} \mathbb{P}(A \mid \{y-\delta \leq Y \leq y+\delta\})$$

$$= \lim_{\delta \to 0} \left(\frac{\mathbb{P}(y-\delta \leq Y \leq y+\delta)\,\mathbb{P}(A)}{\mathbb{P}(y-\delta \leq Y \leq y+\delta)} \right)$$

$$= \lim_{\delta \to 0} \frac{\left(\int_{y-\delta}^{y+\delta} f_{Y|A}(t)\,dt \right) \mathbb{P}(A)}{\int_{y-\delta}^{y+\delta} f_Y(t)\,dt} \cdot \frac{\frac{1}{2\delta}}{\frac{1}{2\delta}}$$

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

After using Total Probability Theorem

$$f_Y(y) = f_{Y|A}(y)\,\mathbb{P}(A) + f_{Y|A^c}(y)\,\mathbb{P}(A^c)$$

$$\mathbb{P}(A \mid \{Y=y\}) = \frac{f_{Y|A}(y)\,\mathbb{P}(A)}{f_{Y|A}(y)\,\mathbb{P}(A) + f_{Y|A^c}(y)\,\mathbb{P}(A^c)}$$

Example - Typical Content For This

- A is an event of the form $\{S=s\}$ - "signal" - S a discrete rv
- Know $\mathbb{P}(\{S=s\})$ $\forall s \in S$
- Y = output of noisy channel you put S through - N
  i.e.
  $$Y = S + N$$
- $N \sim \text{Gaussian}(0, \sigma^2)$
- Observe Y. Want to infer about S

Baye's Rule says that for every s we have

$$\mathbb{P}(\{S=s\}|\{Y=y\}) = \frac{f_{Y|\{S=s\}}(y)\,\mathbb{P}(\{S=s\})}{\sum_{s} f_{Y|\{S=s\}}(y)\,\mathbb{P}(\{S=s\})}$$

Specialize to $S = \pm 1$

$$Y = S + N; \quad N \sim \text{Gaussian}(0, \sigma^2)$$

Suppose

$$\mathbb{P}(\{S=1\}) = p, \quad \mathbb{P}(\{S=-1\}) = 1-p$$

Then

$$f_{Y|\{S=1\}}(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-1)^2}{2\sigma^2}} \sim \text{Gaussian}(1, \sigma^2)$$

finished in Lec21