Hierarchical and mixture models

Introduction (see Section 4.4 of Casella & Berger)

Key idea: joint distributions specified by thinking hierarchically (or conditionally)

$$f(x, y) = f_X(x) f(y|x)$$

Motivations:

1. a model building technique

   (e.g. number of questions answered correctly is Binomial(n, p) but p varies from respondent to respondent)

   $\begin{cases} X|Y \text{ has some dist.} \\ Y \text{ has a dist.} \end{cases} \Longrightarrow$ What is the dist. of X ?

2. Bayesian statistics

$$f(\theta|x) = \frac{f(\theta, x)}{f_X(x)} = \frac{f(x|\theta) f(\theta)}{\int f(x|\theta) f(\theta) d\theta}$$

3. Mixed discrete-continuous models

   (e.g., can take X discrete and $Y|X$ continuous)

Our approach to this topic: consider some examples

Hierarchical and mixture models

Examples

Example 1: pdf $f(x,y) = 1/x$, $0 < y < x < 1$.

We've seen X ~uniform$(0,1)$ and that, given $X = x \in (0,1)$, $Y|X = x$ ~uniform$(0,x)$.

$$f_X(x) = 1 \quad \forall\ 0<x<1, \qquad f(y|x) = \frac{1}{x} \quad \forall\ 0<y<x$$

$$f(x,y) = f(x)f(y|x) = \begin{cases} \frac{1}{x} & 0<x<1,\ 0<y<x \\ 0 & \text{ow} \end{cases}$$

Example 2: Beta-Binomial model

- X =# of male children in n-child family $\quad x = 0,1,2,\dots,n$; $n-x \geq 0$
- P =probability of male child for a random woman
- Consider $X|\mathrm{P} = p$ ~Binominal(n,p)
- Consider P ~Beta(α,β)
- Marginal distribution of X

$X|P=p \sim \text{Bin}(n,p)$

$P \sim \text{Beta}(\alpha,\beta)$

$\Longrightarrow$ what is the dist. of X?

Recall:

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \cap B_i) = \sum \mathbb{P}(A|B_i)\mathbb{P}(B_i)$$

$$\mathbb{P}(A|B_i) = \frac{\mathbb{P}(A \cap B_i)}{\mathbb{P}(B_i)}$$

$$\begin{aligned}
P(X = x) &= \cancel{\mathrm{E}[P(X = x|\mathrm{P})]} \\
&= \int_0^1 \overbrace{P(X = x|\mathrm{P} = p)}^{\text{Bin}(n,p)} f_{\mathrm{P}}(p)dp \\
&= \int_0^1 \binom{n}{x} p^x (1-p)^{n-x} \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)} dp \\
&= \binom{n}{x} \frac{1}{B(\alpha,\beta)} \int_0^1 p^{x+\alpha-1}(1-p)^{n+\beta-x-1}dp \\
&= \binom{n}{x} \frac{B(x+\alpha, n+\beta-x)}{B(\alpha,\beta)}
\end{aligned}$$

$f_{\mathbb{P}}(p)$

$\mathbb{P}(X=x)$

- Expected value of X

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|\mathbb{P})) = \mathbb{E}(n\mathbb{P}) = n\mathbb{E}(\mathbb{P}) = \frac{n\alpha}{\alpha+\beta}$$

Hierarchical models

Examples (cont'd)

Example 3: Poisson-binomial hierarchical model

- Consider $Y|\mathrm{N} \sim$Binominal(N,p)

- Consider N $\sim$Poisson(λ)

- Marginal mean of Y

$$\mathbb{E}(Y) = \mathbb{E}(\underbrace{\mathbb{E}(Y|N)}) = \mathbb{E}(Np) = p\,\mathbb{E}(N) = \lambda p$$

- Marginal distribution of Y

$Y|N \sim Bin(N,p)$

$$\begin{aligned}
P(Y=y) = \sum_{n=0}^{\infty} P(Y=y, N=n) &= \sum_{n=0}^{\infty} P(Y=y|N=n)P(N=n) \\
&= \sum_{n=y}^{\infty} P(Y=y|N=n)P(N=n) \\
&= \sum_{n=y}^{\infty} \binom{n}{y} p^y (1-p)^{n-y} \frac{e^{-\lambda}\lambda^n}{n!} \\
&= \frac{p^y e^{-\lambda}}{y!} \sum_{n=y}^{\infty} (1-p)^{n-y} \frac{\lambda^n}{(n-y)!} \\
& \frac{p^y e^{-\lambda}}{y!} \sum_{z=0}^{\infty} (1-p)^z \frac{\lambda^{z+y}}{z!} \\
&= \frac{(\lambda p)^y e^{-\lambda}}{y!} \sum_{z=0}^{\infty} \frac{[(1-p)\lambda]^z}{z!}
\end{aligned}$$

$Y|N \sim Bin(n,p)$ $\quad N \sim Poisson(\lambda)$

$\mathbb{P}(Y=y|N=n)=0$

$y>n$

$n \geq y$

$n-y := z$

$n = z+y$

$$= \frac{(\lambda p)^y e^{-\lambda}}{y!} e^{(1-p)\lambda}$$

$$= \frac{e^{-\lambda p}(\lambda p)^y}{y!}$$

Recall: $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$

$$\frac{(\lambda p)^y e^{-\lambda}}{y!} \sum_{z=0}^{\infty} \frac{[(1-p)\lambda]^z}{z!} e^{-\lambda(1-p)} \, e^{\lambda(1-p)}$$

$= 1$

$$= e^{\lambda(1-p)} \frac{(\lambda p)^y e^{-\lambda}}{y!} = \frac{e^{-\lambda p}(\lambda p)^y}{y!}$$

$\Rightarrow Y \sim \text{Poisson}(\lambda p)$

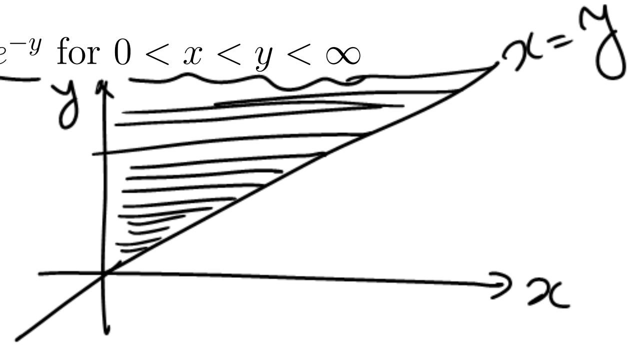
$E(Y) = \lambda p$

Hierarchical and mixture models

Examples (cont'd)

Example 4: continuous X, Y with pdf $f(x, y) = e^{-y}$ for $0 < x < y < \infty$

Where does such a distribution come from?



View 1:

- marginal: $f_X(x) = e^{-x}$ for $0 < x < \infty$
- conditional: $f(y|x) = e^{-(y-x)}$ for $x < y < \infty$
- X is Exponential(1)
- $Y|X = x$ is a location shifted exponential or $(Y - x)|X = x$ is Exponential(1)
- draw an Exponential(1) r.v.to determine X and then, given $X = x$, obtain Y by adding x to another Exponential(1) r.v.

$f(y|x) = \dfrac{f(x,y)}{f(x)}$

$\Rightarrow f(y|x) = \dfrac{e^{-y}}{e^{-x}} = e^{-(y-x)}$

View 2:

- marginal: $f_Y(x) = ye^{-y}$ for $0 < y < \infty$
- conditional: $f(x|y) = y^{-1}$ for $0 < x < y$
- Y is Gamma(2, 1)
- $X|Y = y$ is Uniform$(0, y)$
- given $Y = y$, pick X randomly on $(0, y)$

$f(x|y) = \dfrac{f(x,y)}{f(y)}$