STAT 511 su 2020 Lec 15 slides

Hierarchical models and multinomial distribution

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

- Note that

$$f(y|x) = \frac{f(x,y)}{f_X(x)} \iff f(x,y) = f(y|x)f_X(x).$$

- Often (X, Y) relation is most clearly described by a conditional and marginal.
- A *hierarchical model* describes the joint dist. of an rv pair (X, Y) in the form

$$\begin{aligned} Y|X &\sim \text{Some distribution depending on } X \\ X &\sim \text{Some distribution} \end{aligned}$$

- Can use to get interesting marginal distributions for Y, which take the form

$$f_Y(y) = \int_{-\infty}^{\infty} f(y|x)f_X(x)dx.$$

Poisson-Binomial hierarchical model example

Let

$X = \#$ customers entering a store in a day

$Y = \#$ customers who make purchases

We might assume the following hierarchical model for Y:

$$\begin{aligned} Y|X &\sim \text{Binomial}(X, p) \\ X &\sim \text{Poisson}(\lambda). \end{aligned}$$

Exercise: Find the following:

1. The joint pmf of (X, Y).
2. The marginal pmf of Y.
3. $\mathbb{E}Y$ and $\text{Var}\, Y$.

Theorem (iterated expectation and iterated variance)

For any random variables X and Y we have

- $\mathbb{E}Y = \mathbb{E}(\mathbb{E}[Y|X])$
- $\text{Var } Y = \mathbb{E}(\text{Var}[Y|X]) + \text{Var}(\mathbb{E}[Y|X])$

Exercise: Prove the above.

Exercise: Apply to previous example.

Beta-Binomial hierarchical model example

Let a spectator of a basketball game be chosen at random and let

$$Y = \# \text{ freethrows made out of } n \text{ attempts by the chosen spectator.}$$
$$P = \text{free-throw success rate of the chosen spectator.}$$

We might assume

$$\begin{aligned} Y|P &\sim \text{Binomial}(n, P) \\ P &\sim \text{Beta}(\alpha, \beta). \end{aligned}$$

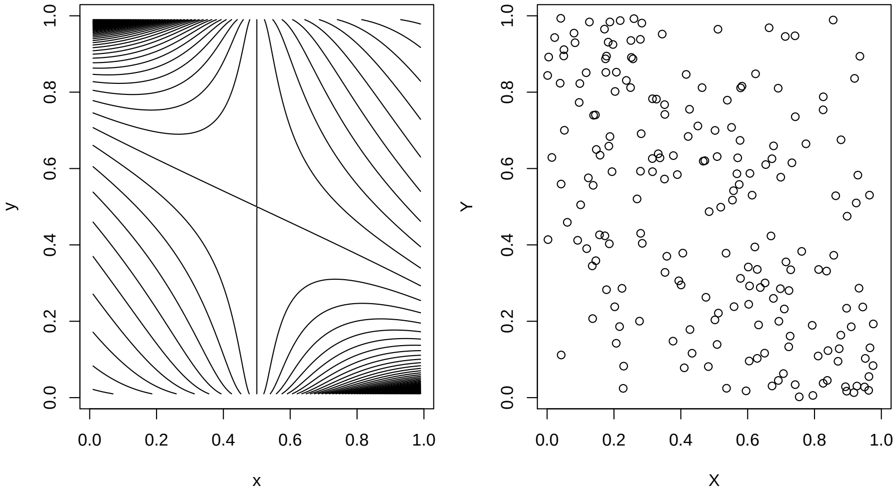
Exercise: Find

1. $\mathbb{E}Y$ and $\text{Var } Y$.
2. The marginal distribution of Y.

Exercise: Let (X, Y) be a pair of rvs such that

$$\begin{aligned} Y|X &\sim \text{Beta}(3/2 - X, 1/2 + X) \\ X &\sim \text{Uniform}(0, 1) \end{aligned}$$

1. Find $\text{Cov}(X, Y)$
2. Find $\text{Var}\, Y$.



0.0
0.2
0.4
0.6
0.8
1.0
x
y
X
Y

Normal random-effects hierarchical model

Let an elementary-school pupil from the U.S. be chosen at random and let

Y be the score on a standardized test the selected pupil.

A be the average test score at the school of the selected pupil.

Suppose we are primarily interested in Y.

We might assume

$$\begin{aligned} Y|A &\sim \text{Normal}(A, \sigma^2) \\ A &\sim \text{Normal}(\mu_A, \sigma_A^2). \end{aligned}$$

Exercise: Find $\mathbb{E}Y$ and $\text{Var}\ Y$.

Multinoulli trial

A *multinoulli trial* is an experiment in which there are K possible outcomes which occur with the probabilities $p_1, \ldots, p_K$, where $\sum_{k=1}^{K} p_k = 1$.

Extends Bernoulli trial (two outcomes) to two or more outcomes.

Multinoulli distribution

Let the random variables $X_1, \dots, X_K$ encode the outcome of a multinoulli trial as

$$X_k = \begin{cases} 1 & \text{if outcome } k \text{ occurs} \\ 0 & \text{otherwise} \end{cases} \quad \text{for } k = 1, \dots, K.$$

Then the set $(X_1, \dots, X_K)$ of K rvs has the *multinoulli distribution* and we write

$$(X_1, \dots, X_K) \sim \text{Multinoulli}(p_1, \dots, p_K).$$

Exercise: Write down the joint pmf of $(X_1, \dots, X_K)$.

Exercise: Find the marginal pmf of X_k for each $k = 1, \dots, K$.

Exercise: Find $\mathbb{E}X_k$ and $\text{Cov}(X_k, X_{k'})$.

Gaussian mixture hierarchical model

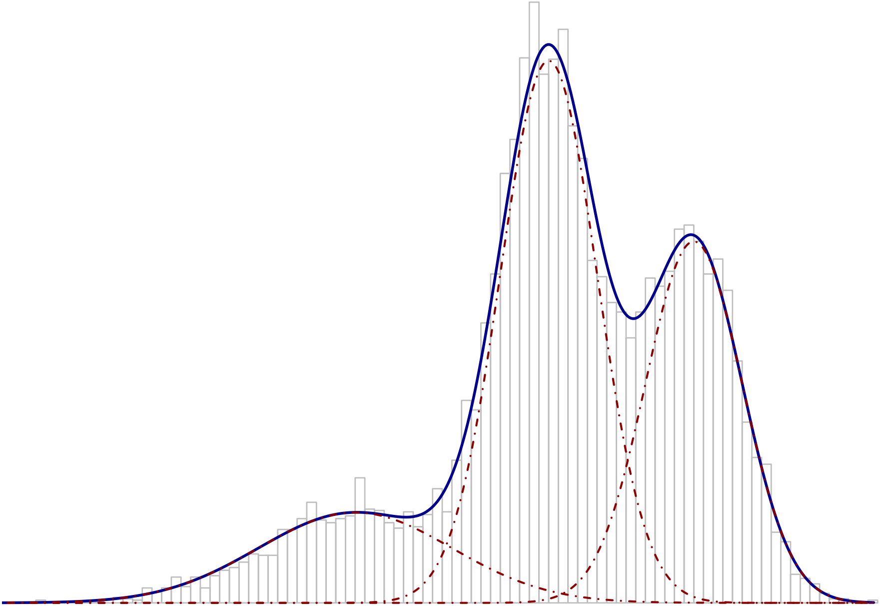
Consider the hierarchical model

$$Y|(X_1,\dots,X_K) \sim \text{Normal}\left(\sum_{k=1}^{K} X_k\mu_k, \sum_{k=1}^{K} X_k\sigma_k^2\right) \quad \text{and}$$

$$(X_1,\dots,X_K) \sim \text{Multinoulli}(p_1,\dots,p_K),$$

where $(\mu_1,\sigma_1^2),\dots,(\mu_K,\sigma_K^2)$ are K mean and variance pairs.

Exercise: Find $\mathbb{E}Y$ and $\text{Var}\, Y$; then find the marginal distribution of Y.



Multinomial distribution

For $k = 1, \dots, K$, let Y_k be # times outcome k occurrs in n independent Multinoulli trials with the outcome probabilities $p_1, \dots, p_K$.

Then the set of rvs $(Y_1, \dots, Y_K)$ has the *multinomial distribution* and we write

$$(Y_1, \dots, Y_K) \sim \text{Multinomial}(n, p_1, \dots, p_K).$$

The joint pmf of $(Y_1, \dots, Y_K)$ is given by

$$p(y_1, \dots, y_K) = \left(\frac{n!}{y_1! \cdots y_K!} \right) p_1^{y_1} \cdots p_K^{y_K}$$

for $(y_1, \dots, y_K) \in \{0, 1, \dots, n\}^K$ st $\sum_{k=1}^K y_k = n$.

Note: For $k = 1, \dots, K$ the marginal distribution of y_k is $\text{Binomial}(n, p_k)$.

Exercise: Suppose each of 10 customers has a coupon for

- ice cream (probability $1/4$)
- French fries (probability $2/4$)
- a donut (probability $1/4$)

The customer's decisions are independent. Let

$$Y_1 = \# \text{ to choose ice cream}$$
$$Y_2 = \# \text{ to choose French fries}$$
$$Y_3 = \# \text{ to choose a donut}$$

Find

1. $P(Y_1 = 5, Y_2 = 3, Y_3 = 2)$
2. $P(Y_2 \geq 4)$
3. $P(Y_1 = 0 | Y_2 = 5)$