Compound Distributions

Definition Suppose an RV, X, follows a certain distribution with parameter θ,

$$X \sim f(x; \theta),$$

but θ itself is another RV with its own distribution,

$$\theta \sim \pi(\theta; \cdot).$$

This effectively turns $f(x; \theta)$ into $f(x|\theta)$, and X now has a so-called compound distribution, with (marginal) distribution

$$f(x) = \sum_{\theta} f(x|\theta)\pi(\theta) \quad \text{or} \quad \int f(x|\theta)\pi(\theta) d\theta.$$

Remark Important concept for hierarchical models.

Example

Setting Suppose $X \sim \text{binomial}(n, p)$.

Question What's the (marginal) distribution of X if

- $n \sim \text{Poisson}(\lambda)$ also random? [Exercise 4.2 (p. 60)]
- $p \sim \text{uniform}(0, 1)$ also random? [Example 3.8 (p. 44–45)]

Context For example,

n = # of hackers who visit your system on a given day;

p = probability of each hacker crashing your system; and

X = total # of crashes they cause.

The "n random" Case

$$f(x) = \sum_{n=0}^{\infty} \underbrace{\left[\frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \right]}_{f(x|n)} \times \underbrace{\left[\frac{e^{-\lambda}\lambda^n}{n!} \right]}_{f(n)}$$

$$\vdots$$

$$= \frac{e^{-p\lambda}(p\lambda)^x}{x!} \quad \sim \quad \text{Poisson}(p\lambda)$$

The “p random” Case

$$f(x) = \int_0^1 \underbrace{\left[\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}\right]}_{f(x|p)} \times \underbrace{(1)}_{f(p)}\, dp$$

$$\vdots$$

$$= \frac{1}{n+1} \quad \sim \quad \text{uniform}\{0, 1, ..., n\}$$

$U \sim$ beta(α, β) if it has density function

$$f(u) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u^{\alpha-1}(1-u)^{\beta-1}, \quad u \in [0,1], \quad \alpha, \beta > 0$$

$\mathbb{E}(\cdot)$ and $\mathbb{V}\text{ar}(\cdot)$

- sometimes, $\mathbb{E}(\cdot)$ and $\mathbb{V}\text{ar}(\cdot)$ are enough, $f(\cdot)$ not necessary
- $\exists$ easy formulae (of course!)
- $f(\cdot)$: law of total probability

$$f(y) = \int f(y|x)f(x)dx \quad \text{or} \quad \sum_x f(y|x)f(x)$$

- $\mathbb{E}(\cdot)$: law of total expectation (aka Adam's Law)†

$$\mathbb{E}(Y) = \mathbb{E}[\mathbb{E}(Y|X)]$$

- $\mathbb{V}\text{ar}(\cdot)$: law of total variance (aka Eve's Law)†

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

---

†The textbook, *Essential Statistics*, does not cover these laws.

Law of Total Expectation

Theorem $\mathbb{E}(Y) = \mathbb{E}[\mathbb{E}(Y|X)]$. [*<u>Note</u>: Outside* $\mathbb{E}[\cdot]$ *w.r.t.* X.]

$$\mathbb{E}(Y) = \int y f(y) dy = \int y \left[\int f(y|x) f(x) dx \right] dy = \ldots$$

$$\ldots = \int \int y f(y|x) f(x) dy dx = \int \left[\int y f(y|x) dy \right] f(x) dx = \ldots$$

$$\ldots = \int \mathbb{E}(Y|X=x) f(x) dx = \mathbb{E}[\mathbb{E}(Y|X)]$$

$$\boxed{\mathbb{E}[g(X)] = \sum_x g(x) f(x) \quad \text{or} \quad \int g(x) f(x) dx}$$

Remark Nick name = Adam's Law.

Mean-Squared-Error Prediction

Theorem Let $g(X) = \mathbb{E}(Y|X)$ and $h(X)$ be any other function of X. Then,

$$\mathbb{E}\{[Y - g(X)]^2\} \le \mathbb{E}\{[Y - h(X)]^2\},$$

i.e., $\mathbb{E}(Y|X)$ is best function of X to predict Y in terms of MSE.

$$\begin{aligned} \text{RHS} &= \mathbb{E}\{[Y - g(X) + g(X) - h(X)]^2\} \\ &= \mathbb{E}\{[Y - g(X)]^2 + [g(X) - h(X)]^2 + 2[Y - g(X)][g(X) - h(X)]\} \\ &= \text{LHS} + (\text{sth} \ge 0) + 2\underbrace{\mathbb{E}\{[Y - g(X)][g(X) - h(X)]\}}_{=0} \end{aligned}$$

Exercise Apply the law of total expectation to show that, if $g(X) = \mathbb{E}(Y|X)$, then $\mathbb{E}\{[Y - g(X)][g(X) - h(X)]\} = 0$.

Law of Total Variance

Theorem $\mathbb{V}\text{ar}(Y) = \mathbb{E}[\mathbb{V}\text{ar}(Y|X)] + \mathbb{V}\text{ar}[\mathbb{E}(Y|X)].$

Intuition (Total $\mathbb{V}$ar) = (Within $\mathbb{V}$ar) + (Between $\mathbb{V}$ar)

x_1	x_2	$\ldots$	x_m		
$y_{1,1}$ $y_{2,1}$ $\vdots$ $y_{n_1,1}$	$y_{1,2}$ $y_{2,2}$ $\vdots$ $y_{n_2,2}$	$\ldots$ $\ldots$ $\vdots$ $\ldots$	$y_{1,m}$ $y_{2,m}$ $\vdots$ $y_{n_m,m}$		
$\mathbb{E}(Y\|x_1)$	$\mathbb{E}(Y\|x_2)$	$\ldots$	$\mathbb{E}(Y\|x_m)$	$\Rightarrow$	$\mathbb{V}\text{ar}[\mathbb{E}(Y\|X)]$
$\mathbb{V}\text{ar}(Y\|x_1)$	$\mathbb{V}\text{ar}(Y\|x_2)$	$\ldots$	$\mathbb{V}\text{ar}(Y\|x_m)$	$\Rightarrow$	$\mathbb{E}[\mathbb{V}\text{ar}(Y\|X)]$

Exercise Prove this law (nick name = Eve's Law).

The "n random" Case

$$\mathbb{E}(X) = \mathbb{E}[\mathbb{E}(X|n)] = ...$$

$$\mathbb{V}\text{ar}(X) = \mathbb{E}[\mathbb{V}\text{ar}(X|n)] + \mathbb{V}\text{ar}[\mathbb{E}(X|n)]$$

$$\vdots$$

	$\mathbb{E}(\cdot)$	$\mathbb{V}\text{ar}(\cdot)$
binomial(n, p)	np	$np(1-p)$
Poisson(λ)	λ	λ

Exercise Verify Adam's Law + Eve's Law in like fashion for the "p random" case.

Random Sums

$$S_N = X_1 + X_2 + \cdots + X_N$$

- $X_1, X_2, \ldots$ are i.i.d. with $\mathbb{E}(X_i) = \mu_x$ and $\mathbb{V}\text{ar}(X_i) = \sigma_x^2$
- N is also random
- e.g., insurance (X_i = individual claim \$; N = # of claims)

$$\begin{aligned} \mathbb{E}(S_N) &= \mathbb{E}[\mathbb{E}(S_N|N)] = \ldots \\ \mathbb{V}\text{ar}(S_N) &= \mathbb{E}[\mathbb{V}\text{ar}(S_N|N)] + \mathbb{V}\text{ar}[\mathbb{E}(S_N|N)] \\ &\vdots \end{aligned}$$

Exercise Compare $\{\mathbb{E}(S_n), \mathbb{V}\text{ar}(S_n)\}$ with $\{\mathbb{E}(S_N), \mathbb{V}\text{ar}(S_N)\}$, where $N \sim \text{Poisson}(n)$.