STAT 801: Mathematical Statistics

Independence, conditional distributions

So far density of X specified explicitly. Often modelling leads to a specification in terms of marginal and conditional distributions.

Def'n: Events A and B are **independent** if

$$P(AB) = P(A)P(B)\,.$$

(Notation: AB is the event that both A and B happen, also written $A \cap B$.)

Def'n: A_i, $i = 1, \ldots, p$ are **independent** if

$$P(A_{i_1} \cdots A_{i_r}) = \prod_{j=1}^{r} P(A_{i_j})$$

for any $1 \le i_1 < \cdots < i_r \le p$.

Example: $p = 3$

$$\begin{aligned} P(A_1A_2A_3) &= P(A_1)P(A_2)P(A_3) \\ P(A_1A_2) &= P(A_1)P(A_2) \\ P(A_1A_3) &= P(A_1)P(A_3) \\ P(A_2A_3) &= P(A_2)P(A_3) \end{aligned}$$

All these equations needed for independence!

Example: Toss a coin twice.

$$\begin{aligned} A_1 &= \{\text{first toss is a Head}\} \\ A_2 &= \{\text{second toss is a Head}\} \\ A_3 &= \{\text{first toss and second toss different}\} \end{aligned}$$

Then $P(A_i) = 1/2$ for each i and for $i \ne j$

$$P(A_i \cap A_j) = \frac{1}{4}$$

but

$$P(A_1 \cap A_2 \cap A_3) = 0 \ne P(A_1)P(A_2)P(A_3)\,.$$

Def'n: X and Y are **independent** if

$$P(X \in A; Y \in B) = P(X \in A)P(Y \in B)$$

for all A and B.

Def'n: Rvs $X_1, \ldots, X_p$ **independent**:

$$P(X_1 \in A_1, \cdots, X_p \in A_p) = \prod P(X_i \in A_i)$$

for any $A_1, \cdots, A_p$.

Theorem:

1. If X and Y are independent then for all x, y
$$F_{X,Y}(x,y) = F_X(x)F_Y(y).$$
2. If X and Y are independent with joint density $f_{X,Y}(x,y)$ then X and Y have densities f_X and f_Y, and
$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$
3. If X and Y independent with marginal densities f_X and f_Y then (X,Y) has joint density
$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$
4. If $F_{X,Y}(x,y) = F_X(x)F_Y(y)$ for **all** x, y then X and Y are independent.
5. If (X,Y) has density $f(x,y)$ and there exist $g(x)$ and $h(y)$ st $f(x,y) = g(x)h(y)$ for (almost) **all** (x,y) then X and Y are independent with densities given by
$$f_X(x) = g(x)/\int_{-\infty}^{\infty} g(u)du$$
$$f_Y(y) = h(y)/\int_{-\infty}^{\infty} h(u)du\,.$$

Proof:

1. Since X and Y are independent so are the events $X \le x$ and $Y \le y$; hence

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

2. Suppose X and Y real valued.

   Asst 2: existence of $f_{X,Y}$ implies that of f_X and f_Y (marginal density formula). Then for any sets A and B

$$\begin{aligned} P(X \in A, Y \in B) &= \int_A \int_B f_{X,Y}(x,y)dydx \\ P(X \in A)P(Y \in B) &= \int_A f_X(x)dx \int_B f_Y(y)dy \\ &= \int_A \int_B f_X(x)f_Y(y)dydx\,. \end{aligned}$$

   Since $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$

$$\int_A \int_B [f_{X,Y}(x,y) - f_X(x)f_Y(y)]dydx = 0\,.$$

   It follows (measure theory) that the quantity in [] is 0 (almost every pair (x, y)).

3. For any A and B we have

$$\begin{aligned} &P(X \in A, Y \in B) \\ &\quad = P(X \in A)P(Y \in B) \\ &\quad = \int_A f_X(x)dx \int_B f_Y(y)dy \\ &\quad = \int_A \int_B f_X(x)f_Y(y)dydx\,. \end{aligned}$$

   If we **define** $g(x, y) = f_X(x)f_Y(y)$ then we have proved that for $C = A \times B$

$$P((X,Y) \in C) = \int_C g(x,y)dydx\,.$$

   To prove that g is $f_{X,Y}$ we need only prove that this integral formula is valid for an arbitrary Borel set C, not just a rectangle $A \times B$.

   This is proved via a *monotone class* argument. The collection of sets C for which identity holds has closure properties which guarantee that this collection includes the Borel sets.

4. Another monotone class argument.

5.

$$\begin{aligned} P(X \in A, Y \in B) &= \int_A \int_B g(x)h(y)dydx \\ &= \int_A g(x)dx \int_B h(y)dy\,. \end{aligned}$$

   Take $B = R^1$ to see that

$$P(X \in A) = c_1 \int_A g(x)dx$$

   where $c_1 = \int h(y)dy$. So $c_1 g$ is the density of X. Since $\int\int f_{X,Y}(xy)dxdy = 1$ we see that $\int g(x)dx \int h(y)dy = 1$ so that $c_1 = 1/\int g(x)dx$. Similar argument for Y.

Theorem: If $X_1,\ldots,X_p$ are independent and $Y_i = g_i(X_i)$ then $Y_1,\ldots,Y_p$ are independent. Moreover, $(X_1,\ldots,X_q)$ and $(X_{q+1},\ldots,X_p)$ are independent.

Conditional probability

Def'n: $P(A|B) = P(AB)/P(B)$ if $P(B) \neq 0$.
Def'n: For discrete X and Y the conditional probability mass function of Y given X is

$$\begin{aligned} f_{Y|X}(y|x) &= P(Y=y|X=x) \\ &= f_{X,Y}(x,y)/f_X(x) \\ &= f_{X,Y}(x,y)/\sum_t f_{X,Y}(x,t) \end{aligned}$$

For absolutely continuous X $P(X=x)=0$ for all x. What is $P(A|X=x)$ or $f_{Y|X}(y|x)$? Solution: use limit

$$P(A|X=x) = \lim_{\delta x \to 0} P(A|x \le X \le x+\delta x)$$

If, e.g., X, Y have joint density $f_{X,Y}$ then with $A = \{Y \le y\}$ we have

$$\begin{aligned} &P(A|x \le X \le x+\delta x) \\ &\qquad = \frac{P(A \cap \{x \le X \le x+\delta x\})}{P(x \le X \le x+\delta x)} \\ &\qquad = \frac{\int_{-\infty}^{y}\int_{x}^{x+\delta x} f_{X,Y}(u,v)\,du\,dv}{\int_{x}^{x+\delta x} f_X(u)\,du} \end{aligned}$$

Divide top, bottom by δx; let $\delta x \to 0$. Denom converges to $f_X(x)$; numerator converges to

$$\int_{-\infty}^{y} f_{X,Y}(x,v)\,dv$$

Define conditional cdf of Y given $X=x$:

$$P(Y \le y|X=x) = \frac{\int_{-\infty}^{y} f_{X,Y}(x,v)\,dv}{f_X(x)}$$

Differentiate wrt y to get def'n of conditional density of Y given $X=x$:

$$f_{Y|X}(y|x) = f_{X,Y}(x,y)/f_X(x)\,;$$

in words "conditional = joint/marginal".