XXVII. Independent Random Variables

Intuition

- We have an experiment with two random variables, Y_1 and Y_2.
- Y_1 and Y_2 are independent if knowing the value of Y_2 gives you no new information about the distribution of Y_1.
- This leads to the formula

$$f_1(y_1) = f(y_1|y_2).$$

Definition and Formulas

- **Definition**: Random variables Y_1 and Y_2 are independent if for all y_1, y_2,

$$P(Y_1 = y_1, Y_2 = y_2) = P(Y_1 = y_1)\, P(Y_2 = y_2).$$

- Short version:

$$\begin{aligned} p(y_1, y_2) &= p_1(y_1)\, p_2(y_2) \quad &\text{(discrete)} \\ f(y_1, y_2) &= f_1(y_1)\, f_2(y_2) \quad &\text{(continuous)} \end{aligned}$$

Note: Since $f(y_1, y_2) = f_2(y_2)\, f(y_1|y_2)$, we could plug this in and cancel $f_2(y_2)$ to get $f_1(y_1) = f(y_1|y_2)$, the formula from the previous slide.

Theorem

- **Theorem**: For continuous random variables, Y_1 and Y_2 are independent if and only if
  1. the domain where $f(y_1, y_2)$ is nonzero is a rectangle (possibly infinite), and
  2. $f(y_1, y_2)$ can be factored into

$$(\text{function of } y_1)(\text{function of } y_2).$$

Example I

Let $f(y_1, y_2) := 6(1 - y_2), 0 \le y_1 \le y_2 \le 1$. Use the definition to determine if Y_1 and Y_2 are independent.

(Graph.)

$$\begin{aligned} f_1(y_1) &= \int_{y_2=y_1}^{y_2=1} 6(1-y_2)\, dy_2 = 3y_1^2 - 6y_1 + 3 \\ f_2(y_2) &= \int_{y_1=0}^{y_1=y_2} 6(1-y_2)\, dy_1 = 6y_2 - 6y_2^2 \end{aligned}$$

We don't have $f(y_1, y_2) = f_1(y_1) f_2(y_2)$, so they're $\boxed{\text{not independent}}$.
This is also clear from the theorem, because the domain is not a rectangle.

Example II

Let $f(y_1, y_2) := y_1 + y_2, 0 \le y_1 \le 1, 0 \le y_2 \le 1$. Use the definition to determine if Y_1 and Y_2 are independent.

$$
\begin{aligned}
f_2(y_2) &:= \int_{y_1=0}^{y_1=1} (y_1 + y_2)\, dy_1 \\
&= \left. \left(\frac{y_1^2}{2} + y_1 y_2 \right) \right|_{y_1=0}^{y_1=1} \\
&= \boxed{y_2 + \frac{1}{2}} \\
f_1(y_1) &= \boxed{y_1 + \frac{1}{2}} \\
f(y_1|y_2) &:= \frac{f(y_1, y_2)}{f_2(y_2)} \\
&= \boxed{\frac{y_1 + y_2}{y_2 + \frac{1}{2}}}
\end{aligned}
$$

$f(y_1, y_2) \neq f_1(y_1) f_2(y_2)$ (or $f_1(y_1) \neq f(y_1|y_2)$), so they're $\boxed{\text{not independent}}$.
Check: This is a rectangle, but it can't be factored, so they're $\boxed{\text{not independent}}$.

Example III

Roll two dice, a red die and a blue die. Define the variables:

$$
\begin{aligned}
Y_1 &:= \text{what shows on red die} \\
Y_2 &:= \text{total}
\end{aligned}
$$

Are Y_1 and Y_2 independent?

Intuitively, no, because if you get a 6 on the red die, then you're more likely to have a high total. We'll check:

$$P(Y_1 = y_1 \cap Y_2 = y_2) = P(Y_1 = y_1)\, P(Y_2 = y_2)\,.$$

Take $y_1 := 6$, $y_2 := 12$. Then

$$\begin{aligned}
P(Y_1 = y_1 \cap Y_2 = y_2) &= \frac{1}{36} \\
P(Y_1 = y_1) &= \frac{1}{6} \\
P(Y_2 = y_2) &= \frac{1}{36} \\
P(Y_1 = y_1 \cap Y_2 = y_2) &\neq P(Y_1 = y_1)\, P(Y_2 = y_2)
\end{aligned}$$

So they're $\boxed{\text{not independent}}$.

Example IV

Consider the joint density function:

$$f(y_1, y_2) := e^{-(y_1+y_2)}, 0 \le y_1 < \infty, 0 \le y_2 < \infty$$

Are Y_1 and Y_2 independent?

$$
\begin{aligned}
f_1(y_1) &:= \int_{y_2=0}^{y_2=\infty} e^{-(y_1+y_2)}\,dy_2 \\
&= e^{-y_1}\int_{y_2=0}^{y_2=\infty} e^{-y_2}\,dy_2 \\
&= e^{-y_1}\left(-e^{-y_2}\right)\Big|_{y_2=0}^{y_2=\infty} \\
&= e^{-y_1} \\
f_2(y_2) &= e^{-y_2} \\
f_1(y_1)\,f_2(y_2) &= e^{-(y_1+y_2)} = f(y_1,y_2)\,\surd
\end{aligned}
$$

So they are $\boxed{\text{independent}}$.
Alternately, we could use the theorem: The region is a rectangle, and $f(y_1, y_2) := e^{-(y_1+y_2)}$ factors into

(function of y_1 only)(function of y_2 only).

Example V

Let $f(y_1, y_2) := 4y_1y_2, 0 \le y_1 \le 1, 0 \le y_2 \le 1$. Are Y_1 and Y_2 independent?

$$\begin{aligned}
f_1(y_1) &:= \int_{y_2=0}^{y_2=\infty} 4y_1y_2\,dy_2 \\
&= e^{-y_1}\ 2y_1y2^2\Big|_{y_2=0}^{y_2=\infty} = 2y_1 \\
f_2(y_2) &= 2y_2 \\
f_1(y_1)\,f_2(y_2) &= 4y_1y_2 = f(y_1, y_2)\ \checkmark
\end{aligned}$$

So they are $\boxed{\text{independent}}$.
Alternately, we could use the theorem: The region is a rectangle, and $f(y_1, y_2) := 4y_1y_2$ factors into

(function of y_1 only)(function of y_2 only).