Joint Distributions

Example: Roll die twice

X = 1st roll
Y = 2nd roll

x \ y	1	2	3	4	5	6	TOTAL
1							$\frac{1}{6}$
2							$\frac{1}{6}$
3							$\frac{1}{6}$
4							$\frac{1}{6}$
5							$\frac{1}{6}$
6							$\frac{1}{6}$
TOTAL							

TWO RANDOM VARIABLES:

Joint PMF

Ex $f_{X,Y}(x,y) = \frac{1}{36}$, $x = 1, \ldots, 6$, $y = 1, \ldots, 6$

→ Can get MARGINAL distributions

$$f_X(x) = \sum_y f(x,y)$$

$$f_Y(y) = \sum_x f(x,y)$$

$f_X(x) = \frac{1}{6}$, $x = 1, \ldots, 6$

$f_Y(y) = \frac{1}{6}$, $y = 1, \ldots, 6$.

Continuous: Joint PDF

$f_{X,Y}(x,y)$

$$\iint f_{X,Y}(x,y)\,dx\,dy = 1$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$

Joint Expectation

(X, Y) → random vector.

$g(X, Y)$ → some function of (X, Y)

$g(x, y) = x + y^2 + \tan(x, y)$.

$$\mathbb{E}\, g(X, Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f_{X,Y}(x,y)\, dx\, dy$$

INDEPENDENCE:

Remember: A, B are indep. events.

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

Independent random variables have joint PMFs / PDFs that satisfy

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$$

for all (x,y) in joint support.

In general, if $X_1, X_2, \ldots, X_n$ are independent then

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdots f_{X_n}(x_n)$$

$$= \prod_{i=1}^{n} f_{X_i}(x_i)$$

Remark: If all n marginal dist'ns above are the same function (say f), then $X_1, \ldots, X_n$ are said to be

→ Independent and Identically Distributed (IID) according to f.

→ a simple random sample (SRS) of size n from the population distribution (f).

Remark: If $\mathbb{X}$ and $\mathbb{Y}$ are independent and u and v are any two functions, then

$$\mathbb{E}\left(u(\mathbb{X})\cdot v(\mathbb{Y})\right)$$

$$=\left(\mathbb{E}u(\mathbb{X})\right)\cdot\left(\mathbb{E}v(\mathbb{Y})\right)$$

Why:

$$\text{LHS} = \iint u(x)\cdot v(y)\, f(x,y)\,dx\,dy$$

$$\left(\int u(x)\, f_{\mathbb{X}}(x)\,dx\right)\left(\int v(y)\, f_{\mathbb{Y}}(y)\,dy\right)$$

BACKGROUND

Given a population, has some distribution, say norm(mean = μ, sd = 7).

Want to learn about μ.

It's difficult to ask everybody in pop'n, so we take SAMPLE instead.

so $\mathbb{X}_1, \ldots, \mathbb{X}_n$ are IID norm(μ, 7).

To estimate (μ), we use $\overline{\mathbb{X}}$.

KNOW: $\mathbb{X}_1, \ldots \mathbb{X}_n$ are RANDOM.

Question: $\overline{\mathbb{X}} \sim$ WHAT ?

Lots of mathematical properties

Proposition: Let X_1 and X_2 be independent with means μ_1, μ_2 and variances σ_1^2, σ_2^2.

Let $Y = aX_1 + bX_2$.

Then

$$EY = E(aX_1 + bX_2)$$

$$= E(aX_1) + E(bX_2)$$

$$= aEX_1 + bEX_2$$

$$EY = a\mu_1 + b\mu_2$$

Variance of Y:

$$\sigma_Y^2 = EY^2 - (EY)^2$$

$$EY^2 = E(aX_1 + bX_2)^2$$

$$= E(a^2X_1^2 + 2abX_1X_2 + b^2X_2^2)$$

$$= E(a^2X_1^2) + E(2abX_1X_2) + Eb^2X_2^2$$

$$= a^2EX_1^2 + 2ab\,EX_1X_2 + b^2EX_2^2$$

$$= a^2(\sigma_1^2 + \mu_1^2) + 2ab\,\mu_1\mu_2 + b^2(\sigma_2^2 + \mu_2^2)$$

$\sigma_1^2 = EX_1^2 - \mu_1^2$, so $EX_1^2 = \sigma_1^2 + \mu_1^2$

$$(EY)^2 = (a\mu_1 + b\mu_2)^2$$

$$= a^2\mu_1^2 + 2ab\,\mu_1\mu_2 + b^2\mu_2^2$$

$$\mathrm{Var}(Y) = a^2\sigma_1^2 + b^2\sigma_2^2$$

Conclusion:

If $T = aX_1 + bX_2$

Then $\mu_T = a\mu_1 + b\mu_2$

$\sigma_T^2 = a^2\sigma_1^2 + b^2\sigma_2^2$

In General: Suppose $X_1, \ldots, X_n$ are ~~a SRS(n)~~ independent with means $\mu_1, \mu_2, \ldots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$

and let $T = a_1X_1 + a_2X_2 + \cdots + a_nX_n$

Then:

$$\mu_T = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n$$

$$= \sum_{i=1}^{n} a_i\mu_i$$

$$\sigma_T^2 = a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2$$

$$= \sum_{i=1}^{n} a_i^2\sigma_i^2$$

HOORAY! If $X_1, \ldots, X_n$ are SRS(n) with common mean $\mu_1 = \mu_2 = \cdots = \mu$

variance $\sigma_1^2 = \cdots = \sigma_n^2 = \sigma^2$

$$\bar{X} = \frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n$$

Then

$$\mu_{\bar{X}} = \mu$$

$$\sigma_{\bar{X}} = \frac{\sigma^2}{n}$$