[Random Vectors and Multivariate Distributions.]

$$(\Omega, \mathcal{A}, P)$$

$\Omega$: Sample Space

$\mathcal{A}$: σ-field, all possible events

$P$: Probability

$$P : \mathcal{A} \to [0,1]$$

$$X : \Omega \to \mathbb{R}.$$

$$\underline{X} : \Omega \to \mathbb{R}^d$$

$$\underline{X}(\omega) = \begin{pmatrix} X_1(\omega) \\ \vdots \\ X_d(\omega) \end{pmatrix}$$

measures d different features of the 'subject' ω.

We will have to deal with $P(X_1 \in A_1, X_2 \in A_2, \ldots X_d \in A_d)$ } joint probabilities

## Midterm 1 details:

Till Section 3.3

<u>Section 2.4</u>: Differentiating under integral sign $\to$ we accomplished this by DCT. So don't worry about the details in this section.

DCT is of course important.

<u>Today: Chapter 4.</u>

d dimensional r. vec:

$$\underline{X} = (X_1, \ldots, X_d) : \Omega \to \mathbb{R}^d$$

such ~~that~~ for every Borel set $\underline{B} \subseteq \mathbb{R}^d$

$$\underline{X}^{-1}(B) \in \underline{\underline{\mathcal{A}}}.$$

i.e $\{\omega : (X_1(\omega), X_2(\omega), \ldots, X_d(\omega)) \in B\} \in \mathcal{A}$

a Borel set of $\mathbb{R}^d$ is simply a set belonging to the Borel $\sigma$-field on $\mathbb{R}^d$,
denoted $\mathcal{B}_{\mathbb{R}^d}$
this being the smallest $\sigma$-field containing all $d$-dimensional open rectangles of the form $(a_1, b_1) \times \cdots \times (a_d, b_d)$

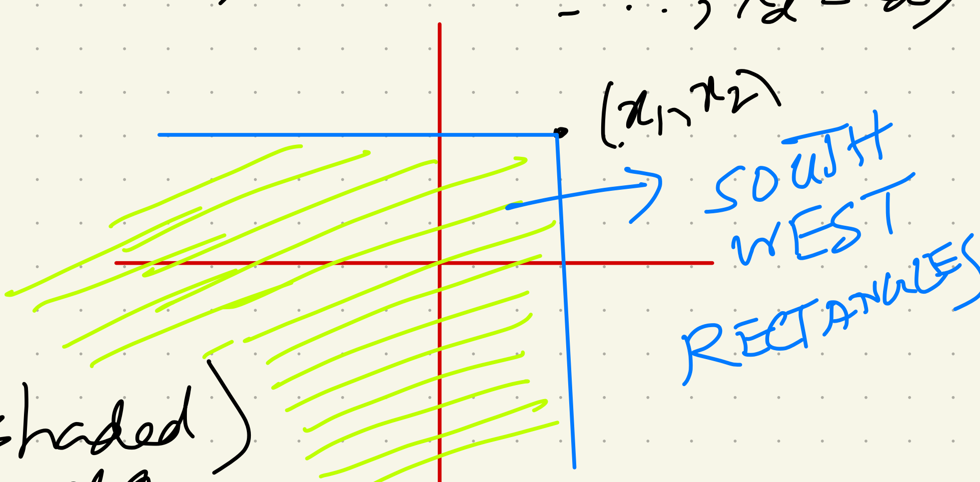
Distributional measure of $\underline{X}$ is

$P_{\underline{X}}$ defined as: $P_{\underline{X}}(\underline{B}) = \underline{P}(\underline{X}^{-1}(B))$
for all $B \in \mathcal{B}_{\mathbb{R}^d}$

Corresponds uniquely to the $d$-dim. dist. fn. of $\underline{X}$ given by: joint distribution function.

$$F_{\underline{X}}(x_1, x_2, \dots, x_n) = P(X_1 \le x_1, X_2 \le x_2, \dots, X_d \le x_d)$$

$d = 2$




$$F_{X_1, X_2}(x_1, x_2) = P((X_1, X_2) \in \text{shaded area})$$

## Simple Example:

Die rolled twice in succession.

$\Omega = \{(i,j) : 1 \le i,j \le 6\}$.

$\mathcal{A}$ = all possible subsets of $\Omega$.

Define $X_1(\omega) = i+j$
$X_2(\omega) = |i-j|$ $\Big\}$ when $\omega = (i,j)$

$P : 2^{\Omega} \longrightarrow [0,1]$.

$P(A)$ = probability ~~that~~ the event $A$ happens.

collection of some $\omega$'s.

$$P(A) = \sum_{(i,j) \in A} \underline{P(\{i,j\})}.$$

$P(\{i,j\}) \geq 0$

how to assign?

$$\sum_{(i,j)} P(\{i,j\}) = 1$$

Infinitely many possible $P$'s on $2^{\Omega}$.

The uniform probability assigns:

$$\left. p(\{i,j\}) = \frac{1}{36} \right\}.$$

$$p(i,j) = \frac{1}{36}$$

$$p : \{1,2,3,4,5,6\}^2 \longrightarrow [0,1].$$

Think about $P_{X_1,X_2}$ under this uniform probability assignment $p$.

$$P_{X_1,X_2}(\{11,0\}) = P[(X_1,X_2) \in \{11,0\}]$$

point in $\mathbb{R}^2$ $\quad$ $B$

$$= P[X_1 = 11,\ X_2 = 0]$$

$$= P[(i,j) : i+j = 11,\ |i-j| = 0]$$

$$\left. = 0 \right\}$$

$P[(X_1,X_2) \in \{m,0\}]$

$= 0$

for any odd $m$.

$S_{X_1, X_2}$ = all pairs in the range of $(X_1, X_2)$ that have non-zero probability under $p_{X_1, X_2}$.

$$\{(k, l) : P(X_1 = k, X_2 = l) > 0\}$$

$$S_{X_1, X_2} = \{(i+j, |i-j|) : 1 \le i, j \le 6\}$$

---

$$S_{X_1, X_2} \subseteq \mathbb{Z}_+ \times \mathbb{Z}_+.$$

$\mathbb{Z}_+$ : all non-negative integers.

---

joint probability mass function:

$$p_{X_1, X_2} : \mathbb{Z}_+ \times \mathbb{Z}_+ \longrightarrow [0, 1]$$

$$p_{X_1, X_2}(k, l) = P(X_1 = k, X_2 = l)$$

Note $p_{X_1, X_2}(k, l) = 0$ for $(k, l) \notin S_{X_1, X_2}$.

$S_{X_1}$ : all $k$ such ~~that~~ $P(X_1 = k) \neq 0$

i.e $p_{X_1}(k) \neq 0$, $p_{X_1}$ : p.m.f of $X_1$.

$$p_{X_1}(8) = P\{(i,j) : i+j = 8\}$$

$$= P\{(2,6), (6,2), (3,5), (5,3), (4,4)\}$$

$$= \frac{5}{36}$$

Conditional distributions: determined by conditional p.m.f.s.

$p_{X_1|X_2=x_2}$ : conditional mass function of $X_1$ given $X_2 = x_2$.

$$\underline{p_{X_1|X_2=x_2}}(x_1) = P(X_1 = x_1 \mid X_2 = x_2)$$

$$= \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_2 = x_2)}$$

only meaningfully defined when $\underbrace{x_2 \in S_{X_2}}_{(p_{X_2}(x_2) > 0)}$.

Check for example $p_{X_1|X_2=2}(8)$

$$= \frac{1}{4}.$$

How do we recover marginal p.m.f's from joint p.m.f?

Take $x_1 \in S_{X_1}$.

$$p_{X_1}(x_1) = P(X_1 = x_1)$$

$$= \sum_{x_2 \in \mathcal{X}_2} \underline{P(X_1 = x_1, X_2 = x_2)}$$

$$= \underline{\sum_{x_2 \in \mathcal{X}_2}} \underbrace{p_{X_1,X_2}(x_1, x_2)}_{\text{could be } 0}$$

$$= \sum_{\substack{x_2 \in \mathcal{X}_2 : \\ (x_1, x_2) \in S_{X_1,X_2}}} p_{X_1,X_2}(x_1, x_2)$$

In general: $S_{X_1} \times S_{X_2} \neq S_{X_1,X_2}$

Example: $11 \in S_{X_1}$, $0 \in S_{X_2}$, $(11, 0) \notin S_{X_1,X_2}$

In general if $(X_1, X_2, \ldots, X_d) \equiv \underline{X}$ is a discrete random vector, its joint pmf:

$$p_{\underline{X}}(x_1, x_2, \ldots, x_d) = P(X_1 = x_1, \ldots, X_d = x_d)$$

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_d \in A_d)$$

$$= \sum_{\substack{(x_1, x_2, \ldots, x_d) \\ \in A_1 \times A_2 \times \cdots \times A_d}} p_{\underline{X}}(x_1, x_2, \ldots, x_d)$$

there are only finitely many d-tuples for which $p_{\underline{X}}(x_1, \ldots, x_d) > 0$.

Independence of $(X_1, \dots, X_d)$: General r.v.s.

We say ~~that~~ $(X_1, \dots, X_d)$ are mutually independent if

$$P_{X_1,\dots,X_d}(\underline{A_1 \times A_2 \times \dots \times A_d})$$
$$= P(X_1 \in A_1, X_2 \in A_1, \dots, X_d \in A_d)$$
$$= \prod_{i=1}^{d} P(X_i \in A_i) = \prod_{i=1}^{d} \underline{P_{X_i}(A_i)}$$

where $A_1, A_2, \dots,$ are Borel subsets of $\mathbb{R}$.

Equivalent to:

$$P(X_1 \in (a_1, b_1), X_2 \in (a_2, b_2), \dots, X_d \in (a_d, b_d))$$
$$= \prod_{i=1}^{d} P(X_i \in (a_i, b_i))$$

$$\Leftrightarrow F_{X_1,\dots,X_d}(x_1, \dots, x_d) = \prod_{i=1}^{d} F_{X_i}(x_i)$$

Further, if $(X_1, \ldots, X_d)$ is a discrete random vector, independence is equivalent to:

$$\boxed{p_{\underline{X}}(x_1, \ldots, x_d) = \prod_{i=1}^{d} p_{X_i}(x_i)}$$

Proof outlined in notes for $d = 2$.

**Multinomial Random Vector:**

Generalization of Binomial Random Variable.

Binomial Rand Var: Multiple replicates of an experiment with 2 possible outcomes.

Now consider: Multiple replicates of an experiment where at each stage experiment results in one of $K$ possible outcomes, $K \geq 2$

EXAMPLE

You're throwing fair die repeatedly.

Outcomes $= \{1, 2, \ldots, 6\}$

$K = 6$.

$p_j$ : the probability of $j^{th}$ outcome.

$$\sum_{j=1}^{K} p_j = 1, \quad p_j > 0.$$

For the dice example, $p_j = \frac{1}{6}$ for $1 \leq j \leq 6$.

We now perform $n$ independent runs of this random expt.

Toss die $n$ times in succession.

No dependence across replications.

$\Omega = \{1, 2, \ldots, K\}$ at any stage

$P(\{j\}) = p_j$

Looking at product space: $\left(\Omega^n, 2^{\Omega^n}, \underline{\underline{P}}^n\right)$

Generic $\omega$ in $\Omega^n$ : a sequence of length $n$ where each slot is one of the symbols 1 through $K$.

$|\Omega^n| = K^n$.

Generic $\omega$ in $\Omega^n = (\eta_1, \eta_2, \ldots, \eta_n)$ where each $\underline{\eta_i}$ is between 1 and $K$.

$$P^n(\{\underline{\underline{\eta_1}}, \underline{\underline{\eta_2}}, \ldots, \underline{\underline{\eta_n}}\}) \longleftarrow$$

will refer to as $P$ henceforth

$$= \prod_{i=1}^{n} \left( \sum_{i} \mathbb{1}(\eta_i = j) \, p_j \right) = \prod_{i=1}^{n} p_{\eta_i}$$

Define: $(N_1, N_2, \ldots, N_K)$ where

$$N_j(\{\underline{\eta_1}, \underline{\eta_2}, \ldots, \underline{\eta_n}\}) = \sum_{i=1}^{n} \mathbb{1}(\eta_i = j)$$

$$\sum_{j=1}^{K} N_j = n.$$ $N_j$ = # of times $j$ appears in the sequence

We say:

$$(N_1, N_2, \ldots, N_K) \sim \text{Mult.}(n, p_1, \ldots, p_K)$$

We'll compute

$$\underline{P(N_1 = n_1, N_2 = n_2, \ldots, N_K = n_K)}$$

$$(n_1, n_2, \ldots, n_K)$$

$$S_{N_1, N_2, \ldots, N_K} = \left\{ (n_1, n_2, \ldots, n_K) : n_i \geq 0, \sum n_i = n \right\}$$

Consider any:

$(\eta_1, \eta_2, \ldots, \eta_n)$ for which

$\underline{N_j = n_j}$

how many sequences are there?

$$P(\{\underline{\eta_1, \eta_2, \ldots, \eta_n}\}) = \prod_{i=1}^{n} p_{\eta_i}$$

$$= \underline{p_1^{n_1} p_2^{n_2} \cdots p_K^{n_K}}$$

Amounts to counting the number of ways $n_1$ 1's, $n_2$ 2's, ..., $n_K$ K's can be permuted.

This is $\dfrac{n!}{n_1!\, n_2! \cdots n_K!}$

$$= \binom{n}{n_1} \times \binom{n-n_1}{n_2} \times \cdots \times \binom{n_K}{n_K}$$

---

So by the law of additivity of probabilities

$$P(N_1 = n_1, \ldots, N_K = n_K)$$

$$= \frac{n!}{\prod_{j=1}^{K} n_j!}\, p_1^{n_1} \cdots p_K^{n_K}$$

Multinomial pmf!