

Def. the covariance of two random variables. Denote as Cov

$$\text{Cov}(X, Y) = E[(X - E X)(Y - E Y)]$$

within some simple calculation, we have:

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

This is a common way to workout the covariance.

and for the calculation of the variance of the sum of numerous random variables.

$$D\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n D(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$$

In particular, we have

$$D(X_1 + X_2) = D(X_1) + D(X_2) + 2\text{Cov}(X_1, X_2)$$

Generally, all the formula above is the special case of the formula:

$$D\left(\sum_{i=1}^n X_i\right) = \sum_{i,j} \text{Cov}(X_i, X_j)$$

where  $\text{Cov}(X_i, X_i) = D(X_i)$

Def. correlation coefficient denote as  $\rho_{ij}$

$$\rho_{ij} = \frac{\text{Cov}(X_i, X_j)}{\sigma_i \sigma_j}$$

There are some properties of correlation coefficient.

1. The absolute value of correlation coefficient is less than or equal to one, i.e.

$$|\rho_{ij}| \leq 1$$

use the Cauchy-Schwarz inequality can solve this.

2.  $\rho$  equals to one iff the two random variables have some kind of linear relation, i.e. exist some  $a, b$  s.t.

$$X_i = aX_j + b$$

$X_i \sim B(n, p_i)$  How we going to work out the correlation coefficient of  $X_i, X_j$ ?

We have some techniques:

Note that  $X_i + X_j \sim B(n, p_i + p_j)$

Consequently

$$E(X_i + X_j) = n(p_i + p_j), D(X_i + X_j) = n(p_i + p_j)(1 - p_i - p_j)$$

$$D(X_i + X_j) = D(X_i) + D(X_j) + 2\text{Cov}(X_i, X_j)$$

It is easy to get that

$$\text{Cov}(X_i, X_j)$$

if  $X, Y$  are 二值函数, then irrelation is equivalent to independence.

Without loss of generality, we consider the 示性函数  $1_A, 1_B$

Apparently

$$\rho = 0 \iff P(AB) - P(A)P(B) = 0$$

that leads to that  $A, B$  are independent.

the normal random variable in multi-dimension space.

Let's put it straight:

$$p(x) \frac{1}{(2\pi)^{n/2} (\det \Sigma)^{1/2}} \exp \left\{ -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right\}$$

Let a finite or infinite sequence of p.m. 's  $\{\mu_j\}$  on  $(\mathcal{R}^1, \mathcal{B}^1)$ , or equivalently their distribution function's be given. There exists a probability space  $(\Omega, \mathcal{F}, \mathcal{P})$  and a sequence of independent random variables  $\{X_j\}$  defined on it such that for each  $j, \mu_j$  is the probability measure of  $X_j$

First, we re-define the language that we're gonna use. The new language includes: (1) proposition:  $A_0, \bar{A}_0, \dots$  Note that for every  $i$ ,  $A_i$  and  $\bar{A}_i$  appear pairwise.

(2) logic operators:  $\vee, \wedge$

(3) parenthesis:  $(, )$

Note: there is some definitions that I don't know how to translate.

Note:  $\neg$  and  $\rightarrow$  is then not the primitive operators.  $\neg\alpha$  is defined as follows.  $\forall A$ , where  $A$  is a proposition,  $\neg A = \bar{A}$ ,  $\neg\bar{A} = A$

$$\neg(\alpha \wedge \beta) = \neg\alpha \vee \neg\beta \text{ and } \neg(\alpha \vee \beta) = \neg\alpha \wedge \neg\beta$$

$$\alpha \rightarrow \beta \equiv \neg\alpha \vee \beta$$

we use  $\Gamma, \alpha$  to denote  $\Gamma \cup \{\alpha\}$ . The rule of prf is described as follows.

Axioms:

$$\Gamma, A_i, \bar{A}_i$$

rule ( $\vee$ ):

$$\frac{\Gamma, \alpha_i}{\Gamma, (\alpha_0 \vee \alpha_1)}$$

rule ( $\wedge$ ):

$$\frac{\Gamma, \alpha_0 \quad \Gamma, \alpha_1}{\Gamma, (\alpha_0 \wedge \alpha_1)}$$

rule of cut:

$$\frac{\Gamma, \alpha_0 \quad \Gamma, \alpha_1}{\Gamma}$$

Example [Ehrenfest chain]. The Ehrenfest chain is a simple model of 'mixing' processes. This chain can shed light on perplexing questions like 'Why aren't people dying all the time due to the air molecules bunching up in some odd corner of their bedrooms while they sleep?'

The model considers  $d$  balls distributed among two urns, and results in a Markov chain  $\{X_0, X_1, \dots\}$  having state space  $\{0, 1, \dots, d\}$ , with the state  $X_n$  of the chain at time  $n$  being the number of balls in urn # 1 at time  $n$ .

At each time , we choose a ball at random uniformly from the  $d$  possibilities, take that ball out of its current urn, and drop it into the other urn. Thus.  $P(i, i-1) = i/d$  and  $P(i, i+1) = (d-i)/d$  for all  $i$

We will soon see a theorem ('a strong law of large numbers for Markov chains') that supports this interpretation.

the identification space.  $B^n/S_{n-1}$  Let  $B^n$  denote the unit ball in  $n$ -dimensional euclidean space, and let  $S^{n-1}$  denote its boundary. Consider the partition of  $B^n$  which has as members:

- (a) the set  $S^{n-1}$
- (b) the individual points of  $B^n - S^{n-1}$

The associated identification space is usually written  $B^n/S^{n-1}$ . In general, if we replace  $B^n$  by an arbitrary space  $X$  and  $S^{n-1}$  by a subspace  $A$ , then  $X/A$  means  $X$  with the subspace  $A$  identified to a point. Note that in this notation,  $CX$  becomes  $X \times I/X \times \{1\}$

We claim that  $B^n/S^{n-1}$  is homeomorphic to  $S^n$ . This is not very surprising. Take for example  $n = 1$ , then we are saying that identifying the endpoints of  $[-1, 1]$  gives a space homeomorphic to a circle. To give a formal proof we need only construct a map  $f : B^n \rightarrow S^n$  which is onto, one-one on  $B^n - S^{n-1}$  and which identifies all of  $S^{n-1}$  to a single point. Our map will be an identification map by corollary (4.4), and so theorem (4.2) provides the required homeomorphism. We can produce  $f$  as follows. We know that  $\mathbb{E}^n$  is homeomorphic to  $B^n - S^{n-1}$  and to  $S^n - \{p\}$  for any point  $p \in S^n$ . Choose specific homeomorphism  $h_1 : B^n - S^{n-1} \rightarrow \mathbb{E}^n$ ,  $h_2 : \mathbb{E}^n \rightarrow S^n - \{p\}$  and define

$$f(x) = \begin{cases} h_2 h_1(x) & \text{for } x \in B^n - S^{n-1} \\ p & \text{for } x \in S^{n-1} \end{cases}$$

the continuity of  $f$  is left to readers to check.

The glueing lemma Let  $X, Y$  be subsets of a topological space and give each of  $X, Y$  and  $X \cup Y$  the induced topology. If  $f : X \rightarrow Z$  and  $g : Y \rightarrow Z$  are functions which agree on the intersection of  $X$  and  $Y$ , we can define

$$f \cup g : X \cup Y \rightarrow Z$$

Comment: pretty cool by  $f \cup g(x) = f(x)$  for  $x \in X$  and  $f \cup g(y) = g(y)$  for  $y \in Y$ . We say that  $f \cup g$  is formed by 'glueing together' the functions  $f$  and  $g$ . The folloing result allows us, under certain conditions, to deduce the continuity of  $f \cup g$  from the continuity of  $f$  and  $g$

Glueing Lemma If  $X$  and  $Y$  are closed in  $X \cup Y$  and if both  $f$  and  $g$  are continuous, then  $f \cup g$  is continuous

Prf. Let  $C$  be a closed subset of  $Z$ . Then  $f^{-1}(C)$  is closed in  $X$ , and therefore closed in  $X \cup Y$  (since  $X$  is closed in  $X \cup Y$ ). Similarly,  $g^{-1}(C)$  is closed in  $X \cup Y$ . But  $(f \cup g)^{-1}(C) = f^{-1}(C) \cup g^{-1}(C)$ , and therefore  $(f \cup g)^{-1}(C)$  is closed in  $X \cup Y$ . This proves  $f \cup g$  is continuous.

The glueing lemma remains true if we ask that  $X$  and  $Y$  are both open in  $X \cup Y$ . We have stated the result for the closed case because it is this case that is most useful in practise. The lemma is of course false if we place no restrictions on  $X$  and  $Y$

As we shall see, the glueing lemma can be explained interms of idenfication maps and interpreted as a special case of theorem (4.3). In order to do this, we introduce the disjoint union  $X + Y$  of the spaces  $X, Y$  and the function  $j : X + Y \rightarrow X \cup Y$  which when restricted to either  $X$  or  $Y$  is just the inclusion in  $X \cup Y$ . This function is important for our purposes because :

- (a) it is continuous
- (b) the composton  $(f \cup g)j : X + Y \rightarrow Z$  is continuous if and only if both  $f$  and  $g$  are continuous.

We have theorem If  $j$  is an identification map, and if both  $f : X \rightarrow X$  and  $g : Y \rightarrow Z$  are continuous, then  $f \cup g : X \cup Y \rightarrow Z$  is continuous.

You are not going to prove this?

The glueing lemma is a special case of this result, since if both  $X$  and  $Y$  are closed in  $X \cup Y$ , then  $j$  sends closed sets to closed sets and is an identification map by theorem 4.3

If  $j$  is an identification map, then we can think of  $X \cup Y$  as an identification space formed from the disjoint union  $X + Y$  by identifying certain points of  $X$  with points of  $Y$ . In this case, we often say that  $X \cup Y$  has the identification topology. The open sets of  $X \cup Y$  are those sets  $A$  for which  $A \cap X$  and  $A \cap Y$  are open in  $X$  and  $Y$  respectively .

$$X_i + \bar{X} = \frac{1}{n} \sum_{j \neq i} X_j + \frac{n+1}{n} X_i$$

that leads to

$$D(X_i + \bar{X}) = \frac{1}{n^2} (n-1) D(X) + \frac{(n+1)^2}{n^2} D(X)$$

$$\frac{n-1+n^2+2n+1}{n^2} D(X) = \frac{n+3}{n} D(X)$$

But

$$D(X_i) + D(\bar{X}) = D(X) + \frac{1}{n} D(X) = \frac{1+n}{n} D(X)$$