

1. THE GOOD AND THE BAD (BUT NOT THE UGLY).

Suppose \mathcal{P} is a set of persons. Let \mathcal{G} be the set of good persons and let \mathcal{B} be the set of bad persons. Let us suppose

$$\mathcal{G} \cup \mathcal{B} = \mathcal{P} \quad \text{and} \quad \mathcal{G} \cap \mathcal{B} = \emptyset.$$

Let

$$p = \frac{|\mathcal{G}|}{|\mathcal{P}|} \quad \text{and let} \quad q = \frac{|\mathcal{B}|}{|\mathcal{P}|};$$

evidently,

$$p \geq 0, \quad q \geq 0 \quad \text{and} \quad p + q = 1.$$

Now fix a positive integer n . Let

$$S = \mathcal{P}^{\{1, \dots, n\}};$$

thus $s \in \Omega$ if

$$s : \{1, \dots, n\} \rightarrow \mathcal{P}.$$

We call $s \in \Omega$ a **sample of \mathcal{P} of size n** . We let S be our sample space and we agree that each sample is equally likely so that

$$P(E) = \frac{|E|}{|S|} \quad \text{whenever } E \subset S.$$

We define the random variables

$$X_1, \dots, X_n$$

by letting

$$X_i(s) = \begin{cases} 1 & \text{if } s(i) \in \mathcal{G}, \\ 0 & \text{if } s(i) \in \mathcal{B}. \end{cases}$$

Then

X_1, \dots, X_n are independent Bernoulli with parameter p .

To see this, for each vector $x = (x_1, \dots, x_n)$ of zeros and ones let

$$O_x = \{i \in \{1, \dots, n\} : x_i = 1\}, \quad \text{let} \quad Z_x = \{i \in \{1, \dots, n\} : x_i = 0\}$$

and let

$$E_x = \{X_1 = x_1, \dots, X_n = x_n\}.$$

Then $s \in E_x$ if and only if

$$s(i) \in \begin{cases} \mathcal{G} & \text{if } i \in O_x, \\ \mathcal{B} & \text{if } i \in Z_x; \end{cases}$$

It follows that, and this is the key point,

$$E_x = \left(|\mathcal{G}|^{|O_x|}\right) \left(|\mathcal{B}|^{|Z_x|}\right)$$

so that, since $n = |O_x| + |Z_x|$,

$$P(E_x) = \frac{\left(|\mathcal{G}|^{|O_x|}\right) \left(|\mathcal{B}|^{|Z_x|}\right)}{|\mathcal{P}|^n} = \left(\frac{|\mathcal{G}|}{|\mathcal{P}|}\right)^{|O_x|} \left(\frac{|\mathcal{B}|}{|\mathcal{P}|}\right)^{n-|O_x|} = p^{|O_x|} q^{n-|O_x|}.$$

Our assertion now follows from stuff we did earlier.