2. Independence and Bernoulli Trials (Euler, Ramanujan and Bernoulli Numbers)

Independence: Events A and B are independent if

$$P(AB) = P(A)P(B). \tag{2-1}$$

- It is easy to show that A, B independent implies $\overline{A}, B$; $A, \overline{B}$; $\overline{A}, \overline{B}$ are all independent pairs. For example, $B = (A \cup \overline{A})B = AB \cup \overline{A}B$ and $AB \cap \overline{A}B = \phi$, so that

$$P(B) = P(AB \cup \overline{A}B) = P(AB) + P(\overline{A}B) = P(A)P(B) + P(\overline{A}B)$$

or

$$P(\overline{A}B) = P(B) - P(A)P(B) = (1 - P(A))P(B) = P(\overline{A})P(B),$$

i.e., $\overline{A}$ and B are independent events.

As an application, let A_p and A_q represent the events

$$A_p = \text{"the prime } p \text{ divides the number } N\text{"}$$

and

$$A_q = \text{"the prime } q \text{ divides the number } N\text{"}.$$

Then from (1-4)

$$P\{A_p\} = \frac{1}{p}, \qquad P\{A_q\} = \frac{1}{q}$$

Also

$$P\{A_p \cap A_q\} = P\{\text{"} pq \text{ divides } N\text{"}\} = \frac{1}{pq} = P\{A_p\}\, P\{A_q\} \tag{2-2}$$

Hence it follows that A_p and A_q are independent events!

- If $P(A) = 0$, then since the event $AB \subset A$ always, we have

$$P(AB) \le P(A) = 0 \Rightarrow P(AB) = 0,$$

and (2-1) is always satisfied. Thus the event of zero probability is independent of every other event!

- Independent events obviously cannot be mutually exclusive, since $P(A) > 0,\ P(B) > 0$ and A, B independent implies $P(AB) > 0.$ Thus if A and B are independent, the event AB cannot be the null set.

- More generally, a family of events $\{A_i\}$ are said to be independent, if for every finite sub collection $A_{i_1}, A_{i_2}, \cdots, A_{i_n},$ we have

$$P\left(\bigcap_{k=1}^{n} A_{i_k}\right) = \prod_{k=1}^{n} P(A_{i_k}). \tag{2-3}$$

• Let

$$A = A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n, \tag{2-4}$$

a union of n independent events. Then by De-Morgan's law

$$\overline{A} = \overline{A}_1 \overline{A}_2 \cdots \overline{A}_n$$

and using their independence

$$P(\overline{A}) = P(\overline{A}_1 \overline{A}_2 \cdots \overline{A}_n) = \prod_{i=1}^{n} P(\overline{A}_i) = \prod_{i=1}^{n} (1 - P(A_i)). \tag{2-5}$$

Thus for any A as in (2-4)

$$P(A) = 1 - P(\overline{A}) = 1 - \prod_{i=1}^{n} (1 - P(A_i)), \tag{2-6}$$

a useful result.

We can use these results to solve an interesting number theory problem.

Example 2.1 Two integers M and N are chosen at random. What is the probability that they are relatively prime to each other?

Solution: Since M and N are chosen at random, whether p divides M or not does not depend on the other number N. Thus we have

$$P\{\text{"}p \text{ divides both } M \text{ and } N\text{"}\}$$

$$= P\{\text{"}p \text{ divides } M\text{"}\}\, P\{\text{"}p \text{ divides } N\text{"}\} = \frac{1}{p^2}$$

where we have used (1-4). Also from (1-10)

$$P\{\text{"}p \text{ does } \textit{not} \text{ divede both } M \text{ and } N\text{"}\}$$

$$= 1 - P\{\text{"}p \text{ divides both } M \text{ and } N\text{"}\} = 1 - \frac{1}{p^2}$$

Observe that "M and N are relatively prime" if and only if there exists no prime p that divides both M and N.

Hence

$$\text{"}M \text{ and } N \text{ are relatively prime"} = \bar{X}_2 \cap \bar{X}_3 \cap \bar{X}_5 \cap \cdots$$

where X_p represents the event

$$X_p = \text{"}p \text{ divides both } M \text{ and } N\text{"}.$$

Hence using (2-2) and (2-5)

$$P\{\text{"}M \text{ and N are relatively prime"}\} = \prod_{p \text{ prime}} P(\bar{X}_p)$$

$$= \prod_{p \text{ prime}} (1 - \tfrac{1}{p^2}) = \frac{1}{\sum_{k=1}^{\infty} 1/k^2} = \frac{1}{\pi^2/6} = \frac{6}{\pi^2} = 0.6079,$$

where we have used the Euler's identity[1]

[1]See Appendix for a proof of Euler's identity by Ramanujan.

$$\sum_{k=1}^{\infty} 1/k^s = \prod_{p \text{ prime}} (1 - \tfrac{1}{p^s})^{-1}.$$

The same argument can be used to compute the probability that an integer chosen at random is "square free".
Since the event

$$\text{"An integer chosen at random is square free"}$$

$$= \bigcap_{p \text{ prime}} \{\text{"}p^2 \text{ does } \textit{not} \text{ divide } N\text{"}\},$$

using (2-5) we have

$$P\{\text{"An integer chosen at random is square free"}\}$$

$$= \prod_{p \text{ prime}} P\{p^2 \text{ does } \textit{not} \text{ divide } N\} = \prod_{p \text{ prime}} (1 - \tfrac{1}{p^2})$$

$$= \frac{1}{\sum_{k=1}^{\infty} 1/k^2} = \frac{1}{\pi^2/6} = \frac{6}{\pi^2}.$$

Note: To add an interesting twist to the 'square free' number problem, Ramanujan has shown through elementary but clever arguments that the inverses of the n^{th} powers of all 'square free' numbers add to S_n / S_{2n}, where (see (2-E))

$$S_n = \sum_{k=1}^{\infty} 1/k^n .$$

Thus the sum of the inverses of the squares of 'square free' numbers is given by

$$\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{5^2}+\frac{1}{6^2}+\frac{1}{7^2}+\frac{1}{10^2}+\frac{1}{11^2}+\frac{1}{13^2}+\frac{1}{14^2}+\cdots=\frac{S_2}{S_4}$$

$$=\frac{\pi^2/6}{\pi^4/90}=\frac{15}{\pi^2}=1.51198.$$

Example 2.2: Three switches connected in parallel operate independently. Each switch remains closed with probability p. (a) Find the probability of receiving an input signal at the output. (b) Find the probability that switch S_1 is open given that an input signal is received at the output.

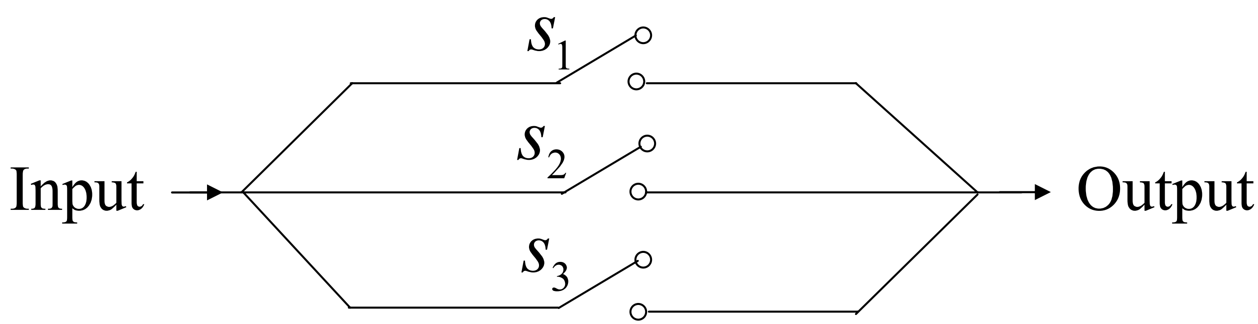

Fig.2.1

Solution: a. Let $A_i =$ "Switch S_i is closed". Then $P(A_i) = p$, $i = 1 \rightarrow 3$. Since switches operate independently, we have

$$P(A_i A_j) = P(A_i)P(A_j);\ \ P(A_1 A_2 A_3) = P(A_1)P(A_2)P(A_3).$$

Let R = "input signal is received at the output". For the event R to occur either switch 1 or switch 2 or switch 3 must remain closed, i.e.,

$$R = A_1 \cup A_2 \cup A_3. \tag{2-7}$$

Using (2-3) - (2-6),

$$P(R) = P(A_1 \cup A_2 \cup A_3) = 1-(1-p)^3 = 3p - 3p^2 + p^3. \tag{2-8}$$

We can also derive (2-8) in a different manner. Since any event and its compliment form a trivial partition, we can always write

$$P(R) = P(R \mid A_1)P(A_1) + P(R \mid \overline{A}_1)P(\overline{A}_1). \tag{2-9}$$

But $P(R \mid A_1) = 1,$ and $P(R \mid \overline{A}_1) = P(A_2 \cup A_3) = 2p - p^2$ and using these in (2-9) we obtain

$$P(R) = p + (2p - p^2)(1-p) = 3p - 3p^2 + p^3, \tag{2-10}$$

which agrees with (2-8).

Note that the events A_1, A_2, A_3 do not form a partition, since they are not mutually exclusive. Obviously any two or all three switches can be closed (or open) simultaneously. Moreover, $P(A_1) + P(A_2) + P(A_3) \neq 1$.

b. We need $P(\overline{A}_1 \mid R)$. From Bayes' theorem

$$P(\overline{A}_1 \mid R) = \frac{P(R \mid \overline{A}_1) P(\overline{A}_1)}{P(R)} = \frac{(2p - p^2)(1-p)}{3p - 3p^2 + p^3} = \frac{2 - 2p + p^2}{3p - 3p^2 + p^3}. \quad (2\text{-}11)$$

Because of the symmetry of the switches, we also have

$$P(\overline{A}_1 \mid R) = P(\overline{A}_2 \mid R) = P(\overline{A}_3 \mid R).$$

Repeated Trials

Consider two independent experiments with associated probability models (Ω_1, F_1, P_1) and (Ω_2, F_2, P_2). Let $\xi \in \Omega_1$, $\eta \in \Omega_2$ represent elementary events. A joint performance of the two experiments produces an elementary events $\omega = (\xi, \eta)$. How to characterize an appropriate probability to this "combined event" ? Towards this, consider the Cartesian product space $\Omega = \Omega_1 \times \Omega_2$ generated from Ω_1 and Ω_2 such that if $\xi \in \Omega_1$ and $\eta \in \Omega_2$, then every ω in Ω is an ordered pair of the form $\omega = (\xi, \eta)$. To arrive at a probability model we need to define the combined trio (Ω, F, P).

Suppose $A \in F_1$ and $B \in F_2$. Then $A \times B$ is the set of all pairs (ξ, η), where $\xi \in A$ and $\eta \in B$. Any such subset of Ω appears to be a legitimate event for the combined experiment. Let F denote the field composed of all such subsets $A \times B$ together with their unions and compliments. In this combined experiment, the probabilities of the events $A \times \Omega_2$ and $\Omega_1 \times B$ are such that

$$P(A \times \Omega_2) = P_1(A), \quad P(\Omega_1 \times B) = P_2(B). \tag{2-12}$$

Moreover, the events $A \times \Omega_2$ and $\Omega_1 \times B$ are independent for any $A \in F_1$ and $B \in F_2$. Since

$$(A \times \Omega_2) \cap (\Omega_1 \times B) = A \times B, \tag{2-13}$$

we conclude using (2-12) that

$$P(A\times B)=P(A\times\Omega_2)\cdot P(\Omega_1\times B)=P_1(A)P_2(B) \tag{2-14}$$

for all $A\in F_1$ and $B\in F_2$. The assignment in (2-14) extends to a unique probability measure $P(\equiv P_1\times P_2)$ on the sets in F and defines the combined trio (Ω, F, P).

Generalization: Given n experiments $\Omega_1,\Omega_2,\cdots,\Omega_n,$ and their associated F_i and $P_i,\ i=1\to n$, let

$$\Omega=\Omega_1\times\Omega_2\times\cdots\times\Omega_n \tag{2-15}$$

represent their Cartesian product whose elementary events are the ordered n-tuples $\xi_1,\xi_2,\cdots,\xi_n,$ where $\xi_i\in\Omega_i$. Events in this combined space are of the form

$$A_1\times A_2\times\cdots\times A_n \tag{2-16}$$

where $A_i\in F_i,$ and their unions an intersections.

If all these n experiments are independent, and $P_i(A_i)$ is the probability of the event A_i in F_i then as before

$$P(A_1 \times A_2 \times \cdots \times A_n) = P_1(A_1)P_2(A_2)\cdots P_n(A_n). \tag{2-17}$$

Example 2.3: An event A has probability p of occurring in a single trial. Find the probability that A occurs exactly k times, $k \leq n$ in n trials.

Solution: Let (Ω, F, P) be the probability model for a single trial. The outcome of n experiments is an n-tuple

$$\omega = \{\xi_1, \xi_2, \cdots, \xi_n\} \in \Omega_0, \tag{2-18}$$

where every $\xi_i \in \Omega$ and $\Omega_0 = \Omega \times \Omega \times \cdots \times \Omega$ as in (2-15). The event A occurs at trial # i, if $\xi_i \in A$. Suppose A occurs exactly k times in ω.

Then k of the ξ_i belong to A, say $\xi_{i_1}, \xi_{i_2}, \cdots, \xi_{i_k}$, and the remaining $n-k$ are contained in its compliment in $\overline{A}$. Using (2-17), the probability of occurrence of such an ω is given by

$$P_0(\omega) = P(\{\xi_{i_1}, \xi_{i_2}, \cdots, \xi_{i_k}, \cdots, \xi_{i_n}\}) = P(\{\xi_{i_1}\})P(\{\xi_{i_2}\})\cdots P(\{\xi_{i_k}\})\cdots P(\{\xi_{i_n}\})$$

$$= \underbrace{P(A)P(A)\cdots P(A)}_{k}\ \underbrace{P(\overline{A})P(\overline{A})\cdots P(\overline{A})}_{n-k} = p^k q^{n-k}. \tag{2-19}$$

However the k occurrences of A can occur in any particular location inside ω. Let $\omega_1, \omega_2, \cdots, \omega_N$ represent all such events in which A occurs exactly k times. Then

$$\text{"}A \text{ occurs exactly } k \text{ times in } n \text{ trials"} = \omega_1 \cup \omega_2 \cup \cdots \cup \omega_N. \tag{2-20}$$

But, all these ω_is are mutually exclusive, and equiprobable.

Thus

$$P(\text{"}A \text{ occurs exactly } k \text{ times in } n \text{ trials"})$$

$$= \sum_{i=1}^{N} P_0(\omega_i) = N P_0(\omega) = N p^k q^{n-k}, \tag{2-21}$$

where we have used (2-19). Recall that, starting with n possible choices, the first object can be chosen n different ways, and for every such choice the second one in $(n-1)$ ways, … and the kth one $(n-k+1)$ ways, and this gives the total choices for k objects out of n to be $n(n-1)\cdots(n-k+1)$. But, this includes the $k!$ choices among the k objects that are indistinguishable for identical objects. As a result

$$N = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{(n-k)!\,k!} \triangleq \binom{n}{k} \tag{2-22}$$

represents the number of combinations, or choices of n identical objects taken k at a time. Using (2-22) in (2-21), we get

$$P_n(k) = P(\text{"}A \text{ occurs exactly } k \text{ times in } n \text{ trials"})$$

$$= \binom{n}{k} p^k q^{n-k}, \quad k = 0,1,2,\cdots,n, \qquad (2\text{-}23)$$

a formula, due to Bernoulli.

Independent repeated experiments of this nature, where the outcome is either a "success" $(= A)$ or a "failure" $(= \overline{A})$ are characterized as Bernoulli trials, and the probability of k successes in n trials is given by (2-23), where p represents the probability of "success" in any one trial.

Example 2.4: Toss a coin n times. Obtain the probability of getting k heads in n trials ?

Solution: We may identify "head" with "success" (A) and let $p = P(H)$. In that case (2-23) gives the desired probability.

Example 2.5: Consider rolling a fair die eight times. Find the probability that either 3 or 4 shows up five times ?

Solution: In this case we can identify

$$\text{"success"} = A = \{ \text{ either } 3 \text{ or } 4 \ \} = \{f_3\} \cup \{f_4\}.$$

Thus

$$P(A) = P(f_3) + P(f_4) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3},$$

and the desired probability is given by (2-23) with $n = 8,\ k = 5$ and $p = 1/3$. Notice that this is similar to a "biased coin" problem.

Bernoulli trial: consists of repeated independent and identical experiments each of which has only two outcomes A or $\overline{A}$ with $P(A) = p$, and $P(\overline{A}) = q$. The probability of exactly k occurrences of A in n such trials is given by (2-23).

Let

$$X_k = \text{"exactly } k \text{ occurrences in } n \text{ trials"}. \tag{2-24}$$

Since the number of occurrences of A in n trials must be an integer $k = 0, 1, 2, \cdots, n$, either X_0 or X_1 or X_2 or $\cdots$ or X_n must occur in such an experiment. Thus

$$P(X_0 \cup X_1 \cup \cdots \cup X_n) = 1. \tag{2-25}$$

But X_i, X_j are mutually exclusive. Thus

$$P(X_0 \cup X_1 \cup \cdots \cup X_n) = \sum_{k=0}^{n} P(X_k) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}. \quad (2\text{-}26)$$

From the relation

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}, \quad (2\text{-}27)$$

(2-26) equals $(p+q)^n = 1,$ and it agrees with (2-25).

For a given *n* and *p* what is the most likely value of *k* ? From Fig.2.2, the most probable value of *k* is that number which maximizes $P_n(k)$ in (2-23). To obtain this value, consider the ratio

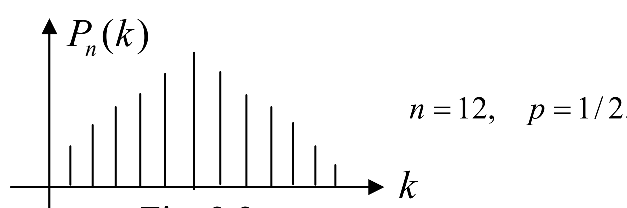

Fig. 2.2

$$\frac{P_n(k-1)}{P_n(k)} = \frac{n!\, p^{k-1} q^{n-k+1}}{(n-k+1)!\,(k-1)!} \frac{(n-k)!\, k!}{n!\, p^k q^{n-k}} = \frac{k}{n-k+1} \frac{q}{p}. \qquad (2\text{-}28)$$

Thus $P_n(k) \geq P_n(k-1)$, if $k(1-p) \leq (n-k+1)p$ or $k \leq (n+1)p$. Thus $P_n(k)$ as a function of *k* increases until

$$k = (n+1)p \qquad (2\text{-}29)$$

if it is an integer, or the largest integer $k_{\max}$ less than $(n+1)p$, and (2-29) represents the most likely number of successes (or heads) in *n* trials.

Example 2.6: In a Bernoulli experiment with *n* trials, find the probability that the number of occurrences of *A* is between k_1 and k_2.

Solution: With $X_i, i = 0,1,2,\cdots,n,$ as defined in (2-24), clearly they are mutually exclusive events. Thus

$$P(\text{" Occurrences of } A \text{ is between } k_1 \text{ and } k_2\text{"})$$

$$= P(X_{k_1} \cup X_{k_1+1} \cup \cdots \cup X_{k_2}) = \sum_{k=k_1}^{k_2} P(X_k) = \sum_{k=k_1}^{k_2} \binom{n}{k} p^k q^{n-k}. \quad (2\text{-}30)$$

Example 2.7: Suppose 5,000 components are ordered. The probability that a part is defective equals 0.1. What is the probability that the total number of defective parts does not exceed 400 ?

Solution: Let

$$Y_k = \text{"} k \text{ parts are defective among 5,000 components"}.$$

Using (2-30), the desired probability is given by

$$P(Y_0 \cup Y_1 \cup \cdots \cup Y_{400}) = \sum_{k=0}^{400} P(Y_k)$$
$$= \sum_{k=0}^{400} \binom{5000}{k} (0.1)^k (0.9)^{5000-k}. \qquad (2\text{-}31)$$

Equation (2-31) has too many terms to compute. Clearly, we need a technique to compute the above term in a more efficient manner.

From (2-29), $k_{\max}$ the most likely number of successes in n trials, satisfy

$$(n+1)p - 1 \le k_{\max} \le (n+1)p \qquad (2\text{-}32)$$

or

$$p - \frac{q}{n} \le \frac{k_{\max}}{n} \le p + \frac{p}{n}, \qquad (2\text{-}33)$$

so that

$$\lim_{n \to \infty} \frac{k_m}{n} = p. \tag{2-34}$$

From (2-34), as $n \to \infty$, the ratio of the most probable number of successes (A) to the total number of trials in a Bernoulli experiment tends to p, the probability of occurrence of A in a single trial. Notice that (2-34) connects the results of an actual experiment (k_m/n) to the axiomatic definition of p. In this context, it is possible to obtain a more general result as follows:

Bernoulli's theorem: Let A denote an event whose probability of occurrence in a single trial is p. If k denotes the number of occurrences of A in n independent trials, then

$$P\left(\left\{ \left| \frac{k}{n} - p \right| > \varepsilon \right\} \right) < \frac{pq}{n\varepsilon^2}. \tag{2-35}$$

Equation (2-35) states that the frequency definition of probability of an event $\frac{k}{n}$ and its axiomatic definition (p) can be made compatible to any degree of accuracy.

Proof: To prove Bernoulli's theorem, we need two identities. Note that with $P_n(k)$ as in (2-23), direct computation gives

$$\begin{aligned}\sum_{k=0}^{n} k\, P_n(k) &= \sum_{k=1}^{n-1} k \frac{n!}{(n-k)!\,k!} p^k q^{n-k} = \sum_{k=1}^{n} \frac{n!}{(n-k)!\,(k-1)!} p^k q^{n-k} \\ &= \sum_{i=0}^{n-1} \frac{n!}{(n-i-1)!\,i!} p^{i+1} q^{n-i-1} = np \sum_{i=0}^{n-1} \frac{(n-1)!}{(n-1-i)!\,i!} p^i q^{n-1-i} \\ &= np(p+q)^{n-1} = np. \end{aligned} \tag{2-36}$$

Proceeding in a similar manner, it can be shown that

$$\begin{aligned}\sum_{k=0}^{n} k^2 P_n(k) &= \sum_{k=1}^{n} k \frac{n!}{(n-k)!\,(k-1)!} p^k q^{n-k} = \sum_{k=2}^{n} \frac{n!}{(n-k)!\,(k-2)!} p^k q^{n-k} \\ &+ \sum_{k=1}^{n} \frac{n!}{(n-k)!\,(k-1)!} p^k q^{n-k} = n^2 p^2 + npq. \end{aligned} \tag{2-37}$$

Returning to (2-35), note that

$$\left|\frac{k}{n}-p\right|>\varepsilon \quad \text{is equivalent to} \quad (k-np)^2>n^2\varepsilon^2, \tag{2-38}$$

which in turn is equivalent to

$$\sum_{k=0}^{n}(k-np)^2 P_n(k)>\sum_{k=0}^{n} n^2\varepsilon^2 P_n(k)=n^2\varepsilon^2. \tag{2-39}$$

Using (2-36)-(2-37), the left side of (2-39) can be expanded to give

$$\sum_{k=0}^{n}(k-np)^2 P_n(k)=\sum_{k=0}^{n}k^2 P_n(k)-2np\sum_{k=0}^{n}k P_n(k)+n^2p^2$$

$$=n^2p^2+npq-2np\cdot np+n^2p^2=npq. \tag{2-40}$$

Alternatively, the left side of (2-39) can be expressed as

$$\sum_{k=0}^{n}(k-np)^2 P_n(k)=\sum_{|k-np|\le n\varepsilon}(k-np)^2 P_n(k)+\sum_{|k-np|>n\varepsilon}(k-np)^2 P_n(k)$$

$$\ge \sum_{|k-np|>n\varepsilon}(k-np)^2 P_n(k)>n^2\varepsilon^2\sum_{|k-np|>n\varepsilon}P_n(k)$$

$$=n^2\varepsilon^2 P\left\{|k-np|>n\varepsilon\right\}. \tag{2-41}$$

Using (2-40) in (2-41), we get the desired result

$$P\left(\left\{\left|\frac{k}{n}-p\right|>\varepsilon\right\}\right)<\frac{pq}{n\varepsilon^2}. \tag{2-42}$$

Note that for a given $\varepsilon > 0,\ pq/n\varepsilon^2$ can be made arbitrarily small by letting n become large. Thus for very large n, we can make the fractional occurrence (relative frequency) $\frac{k}{n}$ of the event A as close to the actual probability p of the event A in a single trial. Thus the theorem states that the probability of event A from the axiomatic framework can be computed from the relative frequency definition quite accurately, provided the number of experiments are large enough. Since $k_{\max}$ is the most likely value of k in n trials, from the above discussion, as $n \to \infty$, the plots of $P_n(k)$ tends to concentrate more and more around $k_{\max}$ in (2-32).

Next we present an example that illustrates the usefulness of "simple textbook examples" to practical problems of interest:

Example 2.8 : **Day-trading strategy :** A box contains n *randomly* numbered balls (not 1 through n but arbitrary numbers including numbers greater than n). Suppose a fraction of those balls $-$ say $m = np \; ; p < 1$ $-$ are initially drawn one by one with replacement while noting the numbers on those balls. The drawing is allowed to continue *until* a ball is drawn with a number larger than the first m numbers. Determine the fraction p to be initially drawn, so as to maximize the probability of drawing the largest among the n numbers using this strategy.

Solution: Let "$X_k = (k+1)^{st}$ drawn ball has the largest number among all n balls, and the largest among the first k balls is in the group of first m balls, $k > m$." (2.43)

Note that X_k is of the form $A \cap B$,

where

A = "largest among the first k balls is in the group of first m balls drawn"

and

B = "$(k+1)^{st}$ ball has the largest number among all n balls".

Notice that A and B are independent events, and hence

$$P(X_k) = P(A)P(B) = \frac{1}{n}\frac{m}{k} = \frac{1}{n}\frac{np}{k} = \frac{p}{k}. \qquad (2\text{-}44)$$

Where $m = np$ represents the fraction of balls to be initially drawn. This gives

P ("selected ball has the largest number among all balls")

$$= \sum_{k=m}^{n-1} P(X_k) = p\sum_{k=m}^{n-1}\frac{1}{k} \approx p\int_{np}^{n}\frac{1}{k} = p\ln k \,\Big|_{np}^{n}$$

$$= -p\ln p. \qquad (2\text{-}45)$$

Maximization of the desired probability in (2-45) with respect to p gives

$$\frac{d}{dp}(-p\ln p) = -(1+\ln p) = 0$$

or

$$p = e^{-1} \simeq 0.3679. \qquad (2\text{-}46)$$

From (2-45), the maximum value for the desired probability of drawing the largest number equals 0.3679 also.
Interestingly the above strategy can be used to "play the stock market".

Suppose one gets into the market and decides to stay up to 100 days. The stock values fluctuate day by day, and the important question is when to get out?

According to the above strategy, one should get out at the first opportunity after 37 days, when the stock value exceeds the maximum among the first 37 days. In that case the probability of hitting the top value over 100 days for the stock is also about 37%. Of course, the above argument assumes that the stock values over the period of interest are randomly fluctuating without exhibiting any other trend. Interestingly, such is the case if we consider shorter time frames such as inter-day trading.

In summary if one must day-trade, then a possible strategy might be to get in at 9.30 AM, and get out any time after 12 noon (9.30 AM + 0.3679×6.5 hrs = 11.54 AM to be precise) at the first peak that exceeds the peak value between 9.30 AM and 12 noon. In that case chances are about 37% that one hits the absolute top value for that day! (disclaimer : Trade at your own risk)

We conclude this lecture with a variation of the *Game of craps* discussed in Example 3-16, Text.

Example 2.9: **Game of craps using biased dice:**

From Example 3.16, Text, the probability of winning the game of craps is 0.492929 $\cdots$ for the player. Thus the game is slightly advantageous to the house. This conclusion of course assumes that the two dice in question are perfect cubes. Suppose that is not the case.

Let us assume that the two dice are slightly loaded in such a manner so that the faces 1, 2 and 3 appear with probability $\frac{1}{6}-\varepsilon$ and faces 4, 5 and 6 appear with probability $\frac{1}{6}+\varepsilon, \quad \varepsilon > 0$ for each dice. If T represents the combined total for the two dice (following Text notation), we get

$$p_4 = P\{T=4\} = P\{(1,3),(2,2),(1,3)\} = 3(\tfrac{1}{6}-\varepsilon)^2$$

$$p_5 = P\{T=5\} = P\{(1,4),(2,3),(3,2),(4,1)\} = 2(\tfrac{1}{36}-\varepsilon^2) + 2(\tfrac{1}{6}-\varepsilon)^2$$

$$p_6 = P\{T=6\} = P\{(1,5),(2,4),(3,3),(4,2),(5,1)\} = 4(\tfrac{1}{36}-\varepsilon^2) + (\tfrac{1}{6}-\varepsilon)^2$$

$$p_7 = P\{T=7\} = P\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} = 6(\tfrac{1}{36}-\varepsilon^2)$$

$$p_8 = P\{T=8\} = P\{(2,6),(3,5),(4,4),(5,3),(6,2)\} = 4(\tfrac{1}{36}-\varepsilon^2) + (\tfrac{1}{6}+\varepsilon)^2$$

$$p_9 = P\{T=9\} = P\{(3,6),(4,5),(5,4),(6,3)\} = 2(\tfrac{1}{36}-\varepsilon^2) + 2(\tfrac{1}{6}+\varepsilon)^2$$

$$p_{10} = P\{T=10\} = P\{(4,6),(5,5),(6,4)\} = 3(\tfrac{1}{6}+\varepsilon)^2$$

$$p_{11} = P\{T=11\} = P\{(5,6),(6,5)\} = 2(\tfrac{1}{6}+\varepsilon)^2.$$

(Note that "(1,3)" above represents the event "the first dice shows face 1, and the second dice shows face 3" etc.)

For $\varepsilon = 0.01,$ we get the following Table:

$T = k$	4	5	6	7	8	9	10	11
$p_k = P\{T = k\}$	0.0706	0.1044	0.1353	0.1661	0.1419	0.1178	0.0936	0.0624

This gives the probability of win on the first throw to be (use (3-56), Text)

$$P_1 = P(T = 7) + P(T = 11) = 0.2285 \qquad (2\text{-}47)$$

and the probability of win by throwing a carry-over to be (use (3-58)-(3-59), Text)

$$P_2 = \sum_{\substack{k=4 \\ k \neq 7}}^{10} \frac{p_k^2}{p_k + p_7} = 0.2717 \qquad (2\text{-}48)$$

Thus

$$P\{\text{winning the game}\} = P_1 + P_2 = 0.5002 \qquad (2\text{-}49)$$

Although perfect dice gives rise to an unfavorable game, a slight loading of the dice turns the fortunes around in favor of the player! (Not an exciting conclusion as far as the casinos are concerned).

Even if we let the two dice to have different loading factors ε_1 and ε_2 (for the situation described above), similar conclusions do follow. For example, $\varepsilon_1 = 0.01$ and $\varepsilon_2 = 0.005$ gives (show this)

$$P\{\text{winning the game}\} = 0.5015. \qquad (2\text{-}50)$$

Once again the game is in favor of the player!

Although the advantage is very modest in each play, from Bernoulli's theorem the cumulative effect can be quite significant when a large number of game are played. All the more reason for the casinos to keep the dice in perfect shape.

In summary, small chance variations in each game of craps can lead to significant counter-intuitive changes when a large number of games are played. What appears to be a favorable game for the house may indeed become an unfavorable game, and when played repeatedly can lead to unpleasant outcomes.

Appendix: Euler's Identity

S. Ramanujan in one of his early papers (*J. of Indian Math Soc*; V, 1913) starts with the clever observation that if $a_2, a_3, a_5, a_7, a_{11}, \cdots$ are numbers less than unity where the subscripts $2,3,5,7,11,\cdots$ are the series of prime numbers, then[1]

$$\frac{1}{1-a_2}\cdot\frac{1}{1-a_3}\cdot\frac{1}{1-a_5}\cdot\frac{1}{1-a_7}\cdots = 1+a_2+a_3+a_2\cdot a_2+a_5$$

$$+a_2\cdot a_3+a_7+a_2\cdot a_2\cdot a_2+a_3\cdot a_3+\cdots. \quad \text{(2-A)}$$

Notice that the terms in (2-A) are arranged in such a way that the product obtained by multiplying the subscripts are the series of all natural numbers $2,3,4,5,6,7,8,9,\cdots$. Clearly, (2-A) follows by observing that the natural numbers

[1]The relation (2-A) is ancient.

are formed by multiplying primes and their powers.

Ramanujan uses (2-A) to derive a variety of interesting identities including the Euler's identity that follows by letting $a_2 = 1/2^s, a_3 = 1/3^s, a_5 = 1/5^s, \cdots$ in (2-A). This gives the Euler identity

$$\prod_{p \text{ prime}} (1 - \tfrac{1}{p^s})^{-1} = \sum_{n=1}^{\infty} 1/n^s. \tag{2-B}$$

The sum on the right side in (2-B) can be related to the Bernoulli numbers (for *s* even).

Bernoulli numbers are positive rational numbers defined through the power series expansion of the even function $\frac{x}{2}\cot(x/2)$. Thus if we write

$$\frac{x}{2}\cot(x/2) \triangleq 1 - B_1 \frac{x^2}{2!} - B_2 \frac{x^4}{4!} - B_3 \frac{x^6}{6!} - \cdots \tag{2-C}$$

then $B_1 = \frac{1}{6}, B_2 = \frac{1}{30}, B_3 = \frac{1}{42}, B_4 = \frac{1}{30}, B_5 = \frac{1}{66}, \cdots.$

By direct manipulation of (2-C) we also obtain

$$\frac{x}{e^x-1}=1-\frac{x}{2}+\frac{B_1x^2}{2!}-\frac{B_2x^4}{4!}+\frac{B_3x^6}{6!}-\cdots \qquad \text{(2-D)}$$

so that the Bernoulli numbers may be defined through (2-D) as well. Further

$$B_n=4n\int_0^\infty \frac{x^{2n-1}}{e^{2\pi x}-1}\,dx=\int_0^\infty x^{2n-1}(e^{-2\pi x}+e^{-4\pi x}+\cdots)dx$$

$$=\frac{2(2n)!}{(2\pi)^{2n}}\left(\frac{1}{1^{2n}}+\frac{1}{2^{2n}}+\frac{1}{3^{2n}}+\frac{1}{4^{2n}}+\cdots\right)$$

which gives

$$S_{2n}\triangleq\sum_{k=1}^{\infty}1/k^{2n}=\frac{(2\pi)^{2n}B_n}{2(2n)!} \qquad \text{(2-E)}$$

Thus[1]

$$\sum_{k=1}^{\infty}1/k^2=\frac{\pi^2}{6};\qquad \sum_{k=1}^{\infty}1/k^4=\frac{\pi^4}{90}\quad \text{etc.}$$

[1]The series $\sum_{k=1}^{\infty} 1/k^2$ can be summed using the Fourier series expansion of a periodic ramp signal as well.