

5

THREE IMPORTANT PROBABILITY DISTRIBUTIONS

10. Bernoulli Trials. The Binomial and Poisson Distributions

By *Bernoulli trials* we mean identical independent experiments in each of which an event A , say, may occur with probability

$$p = \mathbf{P}(A)$$

($p \neq 0$) or fail to occur with probability

$$q = 1 - p.$$

Occurrence of the event A is called a "success," and nonoccurrence of A (i.e., occurrence of the complementary event $\bar{A}$) is called a "failure."

In the case of n consecutive Bernoulli trials, each elementary event ω can be described by a sequence like

$$\underbrace{1011 \dots 0001}_{n \text{ times}}$$

consisting of n digits, each a 0 or a 1, where success at the i th trial is denoted by a 1 in the i th place and failure at the i th trial by a 0 in the i th place. Because of the independence of the trials, the probability of an elementary event ω in which there are precisely k successes and $n - k$ failures is just

$$\mathbf{P}(\omega) =: p^k q^{n-k}.$$

Clearly, the various elementary events are equiprobable only if $p = q$.

Now consider the random variable ξ equal to the total number of successes in n Bernoulli trials, i.e., $\xi(\omega) =: k$ if precisely k successes occur in the

elementary event ω . The number of distinct elementary events with the same total number of successes k is just the number of distinct sequences consisting of k ones and $n - k$ zeros. But the number of such sequences is just the binomial coefficient

$$C_k^n = \binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad (5.1)$$

equal to the number of combinations of n things taken k at a time (recall Theorem 1.3, p. 7). These C_k^n elementary events all have the same probability

$$P(\omega) = p^k q^{n-k},$$

and hence the event $\{\xi = k\}$ has probability

$$P\{\xi = k\} = C_k^n p^k q^{n-k}.$$

Thus the probability distribution of the random variable ξ is given by

$$P_\xi(k) = C_k^n p^k q^{n-k}, \quad k = 0, 1, \dots, n, \quad (5.2)$$

and is known as the *binomial distribution*. The binomial distribution is specified by two parameters, the probability p of a single success and the number of trials n .

It should be noted that the random variable ξ is the sum

$$\xi = \xi_1 + \dots + \xi_n \quad (5.3)$$

of n independent random variables $\xi_1, \dots, \xi_n$, where $\xi_k = 1$ if “success” occurs at the k th trial and $\xi_k = 0$ if “failure” occurs at the k th trial. We have

$$E\xi_k = p, \quad D\xi_k = E\xi_k^2 - (E\xi_k)^2 = p - p^2 = p(1 - p) = pq.$$

Therefore

$$E\xi = np, \quad D\xi = npq. \quad (5.4)$$

Suppose the number of trials is large while the probability of success p is relatively small, so that each success is a rather rare event while the average number of successes np is appreciable. Then it is a good approximation to write

$$P_\xi(k) \sim \frac{a^k}{k!} e^{-a}, \quad k = 0, 1, 2, \dots, \quad (5.5)$$

where $a = np$ is the average number of successes and $e = 2.718\dots$ is the base of the natural logarithms. In fact, we know from calculus that

$$\lim_{n \rightarrow \infty} \left(1 - \frac{a}{n}\right)^n = e^{-a}.$$

But $p = a/n$, and hence (5.2) gives

$$P_{\xi}(0) = q^n = \left(1 - \frac{a}{n}\right)^n \sim e^{-a}.$$

Moreover, it is easily found from (5.1) and (5.2) that

$$\frac{P_{\xi}(k)}{P_{\xi}(k-1)} = \frac{np - (k-1)p}{kq} \sim \frac{a}{k}$$

as $n \rightarrow \infty$. Therefore

$$\begin{aligned} P_{\xi}(1) &\sim \frac{a}{1} P_{\xi}(0) \sim \frac{a}{1} e^{-a}, \\ P_{\xi}(2) &\sim \frac{a}{2} P_{\xi}(1) \sim \frac{a^2}{1 \cdot 2} e^{-a}, \\ &\dots \dots \dots \\ P_{\xi}(k) &\sim \frac{a}{k} P_{\xi}(k-1) \sim \frac{a^k}{k!} e^{-a}, \end{aligned}$$

which proves the approximate formula (5.5).

A random variable ξ taking only the integral values $0, 1, 2, \dots$ is said to have a *Poisson distribution* if

$$P_{\xi}(k) = \frac{a^k}{k!} e^{-a}, \quad k = 0, 1, 2, \dots \quad (5.6)$$

The distribution (5.6) is specified by a single positive parameter a , equal to the mean value of ξ :

$$a = \mathbf{E}\xi = \sum_{k=0}^{\infty} k P_{\xi}(k).$$

In fact, it follows from the expansion

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!},$$

valid for all x , that

$$\mathbf{E}\xi = \sum_{k=0}^{\infty} k P_{\xi}(k) = \sum_{k=0}^{\infty} k \frac{a^k}{k!} e^{-a} = a e^{-a} \sum_{k=1}^{\infty} \frac{a^{k-1}}{(k-1)!} = a e^{-a} e^a = a.$$

Remark. Thus the approximate formula (5.5) shows that the total number of successes in n Bernoulli trials has an approximately Poisson distribution with parameter $a = np$, if n is large and the probability of success p is small.

Example 1 (The lottery ticket problem). How many lottery tickets must be bought to make the probability of winning at least P ?

Solution. Let N be the total number of lottery tickets and M the total number of winning tickets. Then M/N is the probability that a bought ticket is one of the winning tickets. The purchase of each ticket can be regarded as a separate trial with probability of "success" $p = M/N$ in a series of n independent trials, where n is the number of tickets bought. If p is relatively small, as is usually the case, and the given probability P is relatively large, then it is clear that a rather large number of tickets must be bought to make the probability of buying at least one winning ticket no smaller than P . Hence the number of winning tickets among those purchased is a random variable with an approximately Poisson distribution, i.e., the probability that there are precisely k winning tickets among the n purchased tickets is

$$P(k) \approx \frac{a^k}{k!} e^{-a},$$

where

$$a = n \frac{M}{N}.$$

The probability that at least one of the tickets is a winning ticket is just

$$1 - P(0) = 1 - e^{-a}.$$

Hence n must be at least as large as the smallest positive integer satisfying the inequality

$$e^{-a} = e^{-nM/N} \leq 1 - P.$$

Example 2 (The raisin bun problem). Suppose N raisin buns of equal size are baked from a batch of dough into which n raisins have been carefully mixed. Then clearly the number of raisins will vary from bun to bun, although the average number of raisins per bun is just $a = n/N$. What is the probability that any given bun will contain at least one raisin?

Solution. It is natural to assume that the volume of the raisins is much less than that of the dough, so that the raisins move around freely and virtually independently during the mixing, and hence whether or not a given raisin ends up in a given bun does not depend on what happens to the other raisins. Clearly, the raisins will be approximately uniformly distributed throughout the dough after careful mixing, i.e., every raisin has the same probability

$$p = \frac{1}{N}$$

of ending up in a given bun.¹ Imagine the raisins numbered from 1 to n ,

¹ If v is the volume of the raisins and V that of the dough, then $p = v/V$.

and select a bun at random. Then we can interpret the problem in terms of a series of n Bernoulli trials, where "success" at the k th trial means that the k th raisin ends up in the given bun. Suppose both the number of rolls N and the number of raisins n are large, so that in particular $p = 1/N$ is small. Then the number of successes in the n trials, equal to the number of raisins in the given bun, has an approximately Poisson distribution, i.e., the probability $P(k)$ of exactly k raisins appearing in the bun is given by

$$P(k) \approx \frac{a^k}{k!} e^{-a},$$

where

$$a = np = \frac{n}{N}.$$

Hence the probability P of at least one raisin appearing in the bun is

$$P = 1 - P(0) = 1 - e^{-a}.$$

Example 3 (Radioactive decay). It is observed experimentally that radium gradually decays into radon by emitting alpha particles (helium nuclei). The interatomic distances are large enough to justify the assumption that (the nucleus of) each radium atom disintegrates independently of all the others. Moreover, each of the n_0 radium atoms initially present clearly has the same small probability $p(t)$ of disintegrating during an interval of t seconds.² Suppose the disintegration of each radium atom is interpreted as a "success." Then the random variable $\xi(t)$, equal to the number of alpha particles emitted in t seconds, equals the number of successes in a series of n_0 Bernoulli trials with probability of success $p(t)$. The values of n_0 and $p(t)$ are such that the distribution of $\xi(t)$ is very accurately a Poisson distribution, i.e., the probability of exactly k alpha particles being emitted is given by

$$P\{\xi(t) = k\} = \frac{a^k}{k!} e^{-a}, \quad k = 0, 1, 2, \dots, \quad (5.7)$$

where

$$a = E\xi(t) = n_0 p(t)$$

is the average number of alpha particles emitted in t seconds.

Here we have used a model involving Bernoulli trials as a tool for showing that the random variable $\xi(t)$ has a Poisson distribution. Another physical situation leading to a Poisson distribution is considered in Example 4, p. 73.

² A gram of radium ($n_0 \approx 10^{23}$) emits about 10^{10} alpha particles per second. Hence $p(1) \approx 10^{10}/10^{23} = 10^{-13}$.

II. The De Moivre-Laplace Theorem. The Normal Distribution

Next we prove the following basic "limit theorem":

THEOREM 5.1 (De Moivre-Laplace theorem). *Given n independent identically distributed random variables $\xi_1, \dots, \xi_n$, each taking the value 1 with probability p and the value 0 with probability $q = 1 - p$, let*

$$S_n = \sum_{k=1}^n \xi_k, \quad S_n^* = \frac{S_n - ES_n}{\sqrt{DS_n}}.$$

Then

$$\lim_{n \rightarrow \infty} \mathbf{P} \{x' < S_n^* < x''\} = \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} e^{-x^2/2} dx. \quad (5.8)$$

Proof. S_n is the random variable denoted by ξ in (5.3) and (5.4), i.e., S_n is the number of successes in n Bernoulli trials, with mean and variance

$$ES_n = np, \quad DS_n = npq.$$

Hence the "normalized sum" S_n^* is a random variable taking the values

$$x = \frac{k - np}{\sqrt{npq}}, \quad k = 0, 1, \dots, n$$

with probabilities

$$\mathbf{P} \{S_n^* = x\} = P_n(k) = \frac{n!}{k! (n-k)!} p^k q^{n-k}, \quad k = 0, 1, \dots, n.$$

These values divide the interval

$$\left[\frac{-np}{\sqrt{npq}}, \frac{nq}{\sqrt{npq}} \right]$$

into n equal subintervals of length

$$\Delta x = \frac{1}{\sqrt{npq}}.$$

Clearly, as $n \rightarrow \infty$,

$$k = np + \sqrt{npq} x \rightarrow \infty, \quad n - k = nq - \sqrt{npq} x \rightarrow \infty,$$

where the convergence is uniform in every finite interval $x' \leq x \leq x''$.

Using Stirling's formula (see p. 10), we find that

$$\begin{aligned} P_n(k) &\sim \frac{\sqrt{2\pi n} n^n e^{-n}}{\sqrt{2\pi k} k^k e^{-k} \sqrt{2\pi(n-k)} (n-k)^{n-k} e^{-(n-k)}} p^k q^{n-k} \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k(n-k)}} \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k}. \end{aligned}$$

Moreover,

$$\frac{k}{np} = 1 + \sqrt{\frac{q}{np}} x, \quad \frac{n-k}{nq} = 1 - \sqrt{\frac{p}{nq}} x.$$

Therefore, using the expansion

$$\ln(1 + \alpha_n) \sim \alpha_n - \frac{\alpha_n^2}{2}$$

(as $\alpha_n \rightarrow 0$), we have

$$\begin{aligned} \ln \left(\frac{k}{np} \right)^{-k} &= -k \ln \left(1 + \sqrt{\frac{q}{np}} x \right) \\ &\sim -(np + \sqrt{npq} x) \left(\sqrt{\frac{q}{np}} x - \frac{1}{2} \frac{q}{np} x^2 \right), \\ \ln \left(\frac{n-k}{nq} \right)^{-(n-k)} &= -(n-k) \ln \left(1 - \sqrt{\frac{p}{nq}} x \right) \\ &\sim -(nq - \sqrt{npq} x) \left(-\sqrt{\frac{p}{nq}} x - \frac{1}{2} \frac{p}{nq} x^2 \right). \end{aligned}$$

Adding these expressions, we find that

$$\lim_{n \rightarrow \infty} \ln \left(\frac{np}{k} \right)^k \left(\frac{nq}{n-k} \right)^{n-k} = -\frac{x^2}{2},$$

and hence

$$\lim_{n \rightarrow \infty} \left(\frac{np}{k} \right)^k \left(\frac{nq}{n-k} \right)^{n-k} = e^{-x^2/2}$$

uniformly in every finite interval $x' \leq x \leq x''$. Since

$$\sqrt{\frac{n}{n(n-k)}} \sim \sqrt{\frac{k}{np \cdot nq}} = \frac{1}{\sqrt{npq}},$$

it follows that

$$\lim_{n \rightarrow \infty} \mathbf{P} \{S_n^* = x\} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \Delta x, \quad \Delta x = \frac{1}{\sqrt{npq}}.$$

Therefore

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbf{P} \{x' \leq S_n^* \leq x''\} &= \lim_{n \rightarrow \infty} \sum_{x' \leq x \leq x''} \mathbf{P} \{S_n^* = x\} \\ &= \lim_{n \rightarrow \infty} \sum_{x' \leq x \leq x''} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \Delta x, \end{aligned} \quad (5.9)$$

where the sum is over all values of x in the interval $x' < x < x''$. But clearly

$$\lim_{n \rightarrow \infty} \sum_{x' \leq x \leq x''} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \Delta x = \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} e^{-x^2/2} dx \quad (5.10)$$

(why?). Comparing (5.9) and (5.10), we finally get the desired limiting formula (5.8). ■

According to Theorem 5.1, the limiting distribution of the random variable S_n^* is the distribution with probability density

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \quad (5.11)$$

Such a distribution is called a *normal* (or *Gaussian*) *distribution*. The density $p(x)$ is the “bell-shaped” curve shown in Figure 8(a). The corresponding distribution function is

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du, \quad (5.12)$$

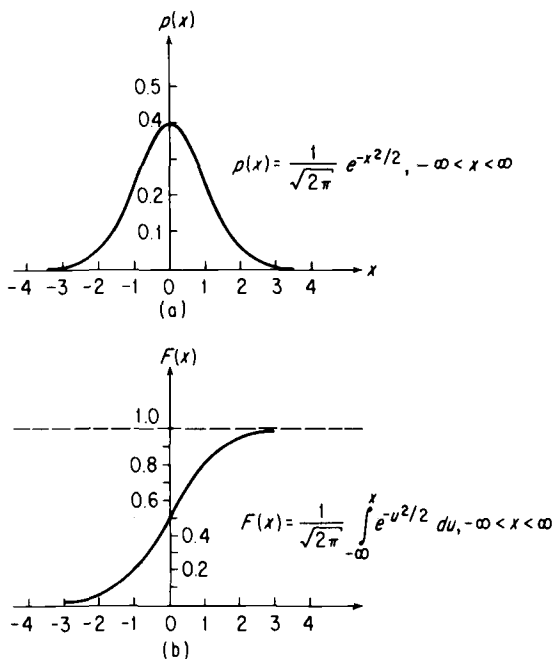


FIGURE 8

Table 2. Values of the normal distribution function $\Phi(x)$ given by formula (5.12).

x	$\Phi(x)$
0.0	0.5000
0.1	0.5398
0.2	0.5793
0.3	0.6179
0.4	0.6554
0.5	0.6915
0.6	0.7257
0.7	0.7580
0.8	0.7881
0.9	0.8159
1.0	0.8413
1.1	0.8643
1.2	0.8849
1.3	0.9032
1.4	0.9192
1.5	0.9332
1.6	0.9452
1.7	0.9554
1.8	0.9641
1.9	0.9713
2.0	0.9773
2.1	0.9821
2.2	0.9861
2.3	0.9893
2.4	0.9918
2.5	0.9938
2.6	0.9953
2.7	0.9965
2.8	0.9974
2.9	0.9981
3.0	0.9986

and is shown in Figure 8(b). Since $p(x)$ is even, it is clear that

$$\Phi(-x) = 1 - \Phi(x).$$

Representative values of $\Phi(x)$ are given in Table 2.

Let ξ be a *normal* (or *Gaussian*) *random variable*, i.e., a random variable with probability density (5.11). Then

$$E\xi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-x^2/2} dx = 0,$$

since the integrand is odd. Moreover,

$$\begin{aligned} D\xi &= E\xi^2 - (E\xi)^2 = E\xi^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \lim_{N \rightarrow \infty} \int_{-N}^N x^2 e^{-x^2/2} dx. \end{aligned}$$

Integrating by parts, we get

$$\begin{aligned} D\xi &= \frac{1}{\sqrt{2\pi}} \lim_{N \rightarrow \infty} \left\{ - \int_{-N}^N x d(e^{-x^2/2}) \right\} \\ &= \frac{1}{\sqrt{2\pi}} \lim_{N \rightarrow \infty} \left\{ \left[-x e^{-x^2/2} \right]_{x=-N}^{x=N} + \int_{-N}^N e^{-x^2/2} dx \right\} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx = 1. \end{aligned}$$

Hence ξ has variance 1. More generally, the random variable with probability density

$$p(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-(x-a)^2/2\sigma^2} \quad (5.13)$$

is also called a normal random variable, and has mean a and variance σ^2 (show this).

Example (Brownian motion). Suppose a tiny particle is suspended in a homogeneous liquid. Then the particle undergoes random collisions with the molecules of the liquid, and, as a result, moves about continually in a chaotic fashion. This is the phenomenon of *Brownian motion*. As a model of Brownian motion, we make the following simplifying assumptions, characterizing a “discrete random walk” in one dimension:

- 1) The particle moves only along the x -axis.
- 2) The particle moves only at the times $t = n\Delta t$, $n = 0, 1, 2, \dots$
- 3) Suppose the particle is at position x at time t . Then, regardless of its previous behavior, the particle moves to either of the two neighboring positions $x + \Delta x$ and $x - \Delta x$ ($\Delta x > 0$) with probability $\frac{1}{2}$.

In other words, at each step the particle undergoes a shift of amount Δx either to the right or to the left, with equal probability.³

Now let $\xi(t)$ denote the position of our "Brownian particle" at time t , and suppose the particle is at the point $x = 0$ at time $t = 0$, so that $\xi(0) = 0$. Then after $t = n\Delta t$ seconds, the particle undergoes n displacements of amount Δx , of which S_n , say, are to the right (the positive direction) and $n - S_n$ to the left (the negative direction). As a result, the position of the particle at time $t = n\Delta t$ is just

$$\xi(t) = [S_n \Delta x - (n - S_n) \Delta x] = (2S_n - n) \Delta x. \quad (5.14)$$

Moreover, since $\xi(0) = 0$, we have

$$\xi(t) = [\xi(s) - \xi(0)] + [\xi(t) - \xi(s)]$$

for any s in the interval $0 < s < t$ (for the time being, s is an integral multiple of Δx). With our assumptions, it is clear that the increments $\xi(s) - \xi(0)$ and $\xi(t) - \xi(s)$ are independent random variables, and that the probability distribution of $\xi(t) - \xi(s)$ is the same as that of $\xi(t - s) - \xi(0)$. Therefore the variance $\sigma^2(t) = D\xi(t)$ satisfies the relation

$$\sigma^2(t) = \sigma^2(s) + \sigma^2(t - s), \quad 0 \leq s \leq t.$$

It follows that $\sigma^2(t)$ is proportional to t , i.e.,⁴

$$D\xi(t) = \sigma^2 t, \quad (5.15)$$

where σ^2 is a constant called the *diffusion coefficient*. On the other hand, it is easy to see that after a time t , i.e., after $n = t/\Delta t$ steps, the variance of the displacement must be

$$D\xi(t) = \frac{t}{\Delta t} (\Delta x)^2. \quad (5.16)$$

Comparing (5.15) and (5.16), we obtain

$$\sigma^2 = \frac{(\Delta x)^2}{\Delta t}. \quad (5.17)$$

The displacements of the particle are independent of one another and can be regarded as Bernoulli trials with probability of "success" $p = \frac{1}{2}$, "success" being interpreted as a displacement in the positive direction. In this sense, the number of displacements S_n in the positive direction is just the number of

³ We will eventually pass to the limit $\Delta t \rightarrow 0$, $\Delta x \rightarrow 0$, thereby getting the "continuous random walk" characteristic of the actual physical process of Brownian motion.

⁴ See footnote 4, p. 40.

successes in n Bernoulli trials. Moreover, the relation between the particle's position at time t and the normalized random variable

$$S_n^* = \frac{S_n - np}{\sqrt{npq}} = \frac{1}{\sqrt{n}} (2S_n - n)$$

is given by

$$\xi(t) = S_n^* \sqrt{n} \Delta x = S_n^* \sqrt{t} \frac{\Delta x}{\sqrt{\Delta t}} = S_n^* \sigma \sqrt{t},$$

because of (5.14) and (5.17). Applying Theorem 5.1, in particular formula (5.8), and passing to the limit $\Delta t \rightarrow 0$ while holding σ constant (so that $\Delta x \rightarrow 0$), we find that the random variable $\xi(t)$ describing the one-dimensional Brownian motion satisfies the formula

$$\mathbf{P} \left\{ x' < \frac{\xi(t)}{\sigma \sqrt{t}} < x'' \right\} = \lim_{\Delta t \rightarrow 0} \mathbf{P} \{ x' < S_n^* < x'' \} = \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} e^{-x^2/2} dx.$$

Therefore $\xi(t)$ is a normal random variable with probability distribution

$$\mathbf{P} \{ x' \leq \xi(t) \leq x'' \} = \frac{1}{\sigma \sqrt{2\pi t}} \int_{x'}^{x''} e^{-x^2/2\sigma^2 t} dx.$$

PROBLEMS

1. Consider the game of "heads or tails," as in Example 3, p. 29. Show that the probability of correctly calling the side of the coin landing upward is always $\frac{1}{2}$ regardless of the call, provided the coin is unbiased. However, show that if the coin is biased, then "heads" should be called all the time if heads are more likely, while "tails" should be called all the time if tails are more likely.

2. There are 10 children in a given family. Assuming that a boy is as likely to be born as a girl, what is the probability of the family having

a) 5 boys and 5 girls; b) From 3 to 7 boys?

3. Suppose the probability of hitting a target with a single shot is 0.001. What is the probability P of hitting the target 2 or more times in 5000 shots?

Ans. $P \approx 1 - 6e^{-5} \approx 0.96$.

4. The page proof of a 500-page book contains 500 misprints. What is the probability P of 2 or more misprints appearing on the same page?

Ans. $P \approx 1 - \frac{5}{2e} \approx 0.08$.

5. Let p be the probability of success in a series of Bernoulli trials. What is the probability P_n of an even number of successes in n trials?

Ans. $P_n = \frac{1}{2} [1 + (1 - 2p)^n]$.

6. What is the probability of the pattern SFS appearing infinitely often in an infinite series of Bernoulli trials, if S denotes "success" and F "failure"?

Hint. Apply the second Borel-Cantelli lemma (Theorem 3.1, p. 33).⁵

Ans. 1.

7. An electronic computer contains 1000 transistors. Suppose each transistor has probability 0.001 of failing in the course of a year of operation. What is the probability of at least 3 transistors failing in a year?

8. A school has 730 students. What is the probability that exactly 4 students were born on January 1?

Hint. Neglect leap years.

9. Let ξ be a random variable with the Poisson distribution (5.6). Find

a) $\sigma^2 = D\xi$; b) $\frac{E(\xi - a)^3}{\sigma^3}$.

Ans. a) a ; b) $\frac{1}{\sqrt{a}}$.

10. Where is the uniform convergence used in the proof of Theorem 5.1?

11. The probability of occurrence of an event A in one trial is 0.3. What is the probability P that the relative frequency of A in 100 independent trials will lie between 0.2 and 0.4?

Hint. Use Theorem 5.1 and Table 2.

Ans. $P \approx 0.97$.

12. Suppose an event A has probability 0.4. How many trials must be performed to assert with probability 0.9 that the relative frequency of A differs from 0.4 by no more than 0.1?

Ans. About 65.

13. The probability of occurrence of an event A in one trial is 0.6. What is the probability P that A occurs in the majority of 60 trials?

Ans. $P \approx 0.94$.

14. Two continuous random variables ξ_1 and ξ_2 are said to have a *bivariate normal distribution* if their joint probability density is

$$p_{\xi_1, \xi_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \times \exp \left\{ -\frac{1}{2(1-r^2)} \left[\frac{(x_1-a)^2}{\sigma_1^2} - 2r \frac{(x_1-a)(x_2-b)}{\sigma_1\sigma_2} + \frac{(x_2-b)^2}{\sigma_2^2} \right] \right\}, \quad (5.18)$$

⁵ For further details, see W. Feller, *op. cit.*, p. 202.

where $\sigma_1 > 0$, $\sigma_2 > 0$, $-1 < r < 1$. Prove that each of the random variables ξ_1 and ξ_2 has a univariate (i.e., one-dimensional) normal distribution of the form (5.13), where $E\xi_1 = a$, $D\xi_1 = \sigma_1^2$, $E\xi_2 = b$, $D\xi_2 = \sigma_2^2$.

Hint. Clearly,

$$p_{\xi_1}(x_1) = \int_{-\infty}^{\infty} p_{\xi_1, \xi_2}(x_1, x_2) dx_2, \quad p_{\xi_2}(x_2) = \int_{-\infty}^{\infty} p_{\xi_1, \xi_2}(x_1, x_2) dx_1$$

(why?).

15. Prove that the number r in (5.18) is the correlation coefficient of the random variables ξ_1 and ξ_2 . Prove that ξ_1 and ξ_2 are independent if and only if $r = 0$.

Comment. This is a situation in which r is a satisfactory measure of the extent to which the random variables ξ_1 and ξ_2 are dependent (the larger $|r|$, the “more dependent” ξ_1 and ξ_2).

16. Let ξ_1 and ξ_2 be the same as in Problem 14. Find the probability distribution of $\eta = \xi_1 + \xi_2$.

Ans. The random variable η is normal, with probability density

$$p_{\eta}(x) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + 2r\sigma_1\sigma_2 + \sigma_2^2)}} \exp \left[-\frac{(x - a - b)^2}{2(\sigma_1^2 + 2r\sigma_1\sigma_2 + \sigma_2^2)} \right].$$