where n is the total frequency, i.e. the total number of sets of m trials each.

Example 8.1 Twelve dice were thrown 2,630 times and each time the number of dice which had 5 or 6 on the uppermost face was recorded. The results are shown in the following table :

Number of dice with 5 or 6 uppermost	0	1	2	3	4	5	6	7	8	9	10	11	12
Frequency	18	115	326	548	611	519	307	133	40	11	2	0	0

Graduate the observed distribution (a) with a binomial distribution for which p is unknown and (b) with a binomial distribution for which $p = \frac{1}{3}$.

Case 1 : Here to fit a binomial distribution, p has to be estimated from the observed distribution. The mean of the latter distribution is

$$\bar{x} = \frac{\sum_x xf_x}{n} = \frac{10{,}662}{2{,}630} = 4 \cdot 05399\,;$$

so the estimate of p is

$$\hat{p} = \frac{4 \cdot 05399}{12} = 0 \cdot 33783\,.$$

The probabilities $f(x)$ are calculated by using the relation

$$f(x) = \left(\frac{m-x+1}{x} \times \frac{\hat{p}}{\hat{q}}\right) \times f(x-1),$$

for $x = 1,\ 2,\ \ldots\ldots,\ m$. Here

$$f(0) = \hat{q}^m,$$

or

$$\log f(0) = m \log \hat{q} = 12 \times \log 0 \cdot 66217$$

$$= \bar{3} \cdot 8516340 = \log 0 \cdot 0071061\,,$$

so that $f(0) = 0{\cdot}0071061.$

Also, $\hat{p}/\hat{q} = 0 \cdot 51019\,.$

The subsequent calculations are shown in Table 8.2 :

A comparison of the last two columns of the table indicates that the fit has been quite satisfactory.

TABLE 8.2

FITTING A BINOMIAL DISTRIBUTION TO THE FREQUENCY DISTRIBUTION OF NUMBER OF DICE SHOWING 5 OR 6 IN 2,630 THROWS OF 12 DICE (p ESTIMATED FROM DATA)

x (1)	$\frac{m-x+1}{x}$ (2)	col. (2) × $\hat{p}/\hat{q}$ (3)	$f(x) = f(x-1)$ × col. (3) (4)	Expected frequency $= n \times$ col. (4) (5)	Observed frequency (6)
0	—	—	0·0071061	18·69	18
1	12	6·12228	0·0435055	114·42	115
2	5·5	2·80604	0·1220782	321·07	326
3	3·33333	1·70063	0·2076098	546·01	548
4	2·25	1·14793	0·2383215	626·79	611
5	1·6	0·81630	0·1945418	511·64	519
6	1·16667	0·59522	0·1157952	304·54	307
7	0·85714	0·43730	0·0506372	133·18	133
8	0·625	0·31887	0·0161467	42·47	40
9	0·44444	0·22675	0·0036613	9·63	11
10	0·3	0·15306	0·0005604	1·47	2
11, 12	—	—	0·0000363*	0·09	0
Total	—	—	1·0000000	2,630·00	2,630

*Obtained from the identity : $f(11) + f(12) = 1 - \sum_{x=0}^{10} f(x)$

Case 2 : Here the procedure is the same as in Case 1, but for p we now use its given value, 1/3. So

$$f(0) = q^m = (2/3)^{12}.$$

or

$$\log f(0) = 12(\log 2 - \log 3)$$
$$= \overline{3}\cdot 8869044 = \log 0\cdot 0077073,$$

giving

$$f(0) = 0\cdot 0077073.$$

We also note that

$$p/q = \tfrac{1}{2}.$$

The expected frequencies may then be calculated as in Case 1.

Here also a comparison of cols. (5) and (6) of the Table 8.3 shows that the fit has been fairly satisfactory, although it is less good than in Case 1.

TABLE 8.3

FITTING A BINOMIAL DISTRIBUTION TO THE FREQUENCY DISTRIBUTION OF NUMBER OF DICE SHOWING 5 OR 6 IN 2,630 THROWS OF 12 DICE ($p = \frac{1}{3}$)

x (1)	$\frac{m-x+1}{x}$ (2)	col. (2) × p/q (3)	$f(x) = f(x-1)$ × col. (3) (4)	Expected frequency = n × col. (4) (5)	Observed frequency (6)
0	—	—	0·0077073	20·27	18
1	12	6·00000	0·0462438	121·62	115
2	5·5	2·75000	0·1271704	334·46	326
3	3·33333	1·66667	0·2119511	557·43	548
4	2·25	1·12500	0·2384450	627·11	611
5	1·6	0·80000	0·1907560	501·69	519
6	1·16667	0·58333	0·7112737	292·65	307
7	0·85714	0·42857	0·0476886	125·42	133
8	0·625	0·31250	0·0149027	39·19	40
9	0·44444	0·22222	0·0033117	8·71	11
10	0·3	0·15000	0·0004968	1·31	2
11, 12	—	—	0·0000529*	0·14	0
Total	—	—	1·0000000	2,630·00	2,630

*Obtained from the identity : $f(11) + f(12) = 1 - \sum_{x=0}^{10} f(x)$.

8.10 The Poisson distribution

This, again, is a discrete distribution. It has the p.m.f.

$$f(x) = \exp(-\lambda)\, \lambda^x/x! \quad \text{if } x = 0, 1, 2, \ldots\ldots$$
$$= 0 \quad \text{otherwise,} \qquad \ldots (8.40)$$

where the parameter λ is a positive quantity.

It is readily seen that $f(x) \geq 0$ for all x. Further,

$$\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} \exp(-\lambda)\lambda^x/x! = \exp(-\lambda)\sum_{x=0}^{\infty} \lambda^x/x!$$
$$= \exp(-\lambda)\ \exp(\lambda) = 1.$$

The distribution may be looked upon as a limiting form of the binomial distribution. Thus suppose for a binomial distribution, $m \to \infty$

8.12 A recursion relation for moments of the Poisson distribution

Like the moments of the binomial distribution, the moments of the Poisson distribution are linked by a recursion relation.

For the Poisson distribution, we have

$$\mu_r = \sum_{x=0}^{\infty}(x-\lambda)^r \exp(-\lambda)\frac{\lambda^x}{x!}.$$

Differentiating both sides with respect to λ, we get

$$\frac{d\mu_r}{d\lambda} = -r\sum_{x=0}^{\infty}(x-\lambda)^{r-1}\exp(-\lambda)\frac{\lambda^x}{x!} - \sum_{x=0}^{\infty}(x-\lambda)^r \exp(-\lambda)\frac{\lambda^x}{x!}$$

$$+\sum_{x=0}^{\infty}(x-\lambda)^r \exp(-\lambda)\frac{x\lambda^{x-1}}{x!}$$

$$= -r\mu_{r-1} + \sum_{x=0}^{\infty}(x-\lambda)^r \frac{\exp(-\lambda)\lambda^{x-1}}{x!}(x-\lambda)$$

or
$$\lambda\left(r\mu_{r-1} + \frac{d\mu_r}{d\lambda}\right) = \sum_{x=0}^{\infty}(x-\lambda)^{r+1}\frac{\exp(-\lambda)\lambda^x}{x!}$$

or
$$\mu_{r+1} = \lambda\left(r\mu_{r-1} + \frac{d\mu_r}{d\lambda}\right). \qquad \ldots (8.49)$$

Putting $\mu_0 = 1$ and $\mu_1 = 0$ in (8.49) and taking $r = 1, 2, 3$, etc., successively, one can obtain the central moments of higher orders.

8.13 Fitting a Poisson distribution to an observed distribution

The Poisson distribution has only one parameter, viz. λ, which can be estimated from the observed data by the method of moments. The mean of the Poisson distribution is λ, while the mean of the observed distribution is $\bar{x}$. The method of moments, therefore, requires that we take as our estimate

$$\hat{\lambda} = \bar{x}.$$

The expected frequencies corresponding to the observed frequencies will then be obtained as

$$n \times f(x) = n \times \frac{\exp(-\hat{\lambda})\hat{\lambda}^x}{x!}, \text{ for } x = 0, 1, 2, \text{ etc.} \qquad \ldots (8.50)$$

Example 8.2 The following table gives the frequency distribution of number of weed seeds per packet for 196 one-lb. packets of a variety of pulses :

Number of weed seeds	0	1	2	3	4	5	6	7	8	9
Frequency	7	33	54	37	34	16	8	5	1	1

In order to fit a Poisson distribution to these data, we first have, as an estimate $\hat{\lambda}$ of the parameter λ, the observed mean

$$\bar{x} = \sum_x x f_x / n$$
$$= 568/196 = 2{\cdot}898.$$

The expected frequencies are then obtained from the formula

$$n \times f(x) = 196 \times \frac{\exp[-2{\cdot}898](2{\cdot}898)^x}{x!}.$$

We note that

$$\log \hat{\lambda} = 0{\cdot}4620983$$

TABLE 8.5

FITTING A POISSON DISTRIBUTION TO THE FREQUENCY DISTRIBUTION OF NUMBER OF WEED SEEDS PER PACKET OF PULSES

Number of weed seeds per packet x (1)	$x!$ (2)	$x \log \hat{\lambda} - \hat{\lambda} \log e$ (3)	$f(x)$ = antilog (3) / (2) (4)	Expected frequency $= n \times (4)$ (5)	Observed frequency (6)
0	1	$\bar{2}{\cdot}7414145$	0·055133	10·79	7
1	1	$\bar{1}{\cdot}2035128$	0·15978	31·32	33
2	2	$\bar{1}{\cdot}6656111$	0·23152	45·38	54
3	6	0·1277094	0·22364	43·83	37
4	24	0·5898077	0·16203	31·76	34
5	120	1·0519060	0·093913	18·41	16
6	720	1·5140043	0·045360	8·89	8
7	5040	1·9761026	0·018779	3·68	5
8	40320	2·4382009	0·006803	1·33	1
9	362880	2·9002992	0·002190	0·43	1
≥ 10	—	—	—	0·18*	0
Total	—	—	—	196·00	196

*Obtained from the identity : $\sum_{x=10}^{\infty} f(x) = 1 - \sum_{x=0}^{9} f(x)$.

and $$\log \exp(-\hat{\lambda}) = -\hat{\lambda} \log e$$
$$= -2{\cdot}898 \times 0{\cdot}4342945$$
$$= -1{\cdot}2585855.$$

The subsequent calculations are shown in Table 8.5.

(As the reader may well realise, one can obtain the value of $f(0)$ as antilog $(-1{\cdot}2585855) = 0{\cdot}055133$ and then calculate

$$f(1) = \frac{\hat{\lambda}}{1} \times f(0),\ f(2) = \frac{\hat{\lambda}}{2} \times f(1),\ f(3) = \frac{\hat{\lambda}}{3} \times f(2),$$

and so on. The defect of this alternative method is that here the errors of approximation accumulate and so probabilities for values towards the end of the table may be rendered highly unreliable.)

Comparing the last two columns of this table, we may say that the fit has been only fairly good.

8.14 The negative binomial distribution

This is another distribution of the discrete type. It has the probability-mass function

$$f(x) = \binom{r+x-1}{r-1} p^r q^x \quad \text{if } x = 0, 1, 2, \ldots\ldots$$
$$= 0 \qquad \text{otherwise}, \qquad \ldots \quad (8.51)$$

where r is a positive integer, $0 < p < 1$ and $q = 1 - p$.

In case $r = 1$, i.e. in case

$$f(x) = pq^x \qquad \text{if } x = 0, 1, 2, \ldots\ldots$$
$$= 0 \qquad \text{otherwise},$$

we get what is said to be a *geometric* distribution.

Note that $f(x) \geq 0$ for all x,

and $$\sum_{x=0}^{\infty} f(x) = p^r \sum_{x=0}^{\infty} \binom{r+x-1}{r-1} q^x$$
$$= p^r \sum_{x=0}^{\infty} \binom{r+x-1}{x} q^x$$
$$= p^r (1-q)^{-r} = 1,$$

so that $f(x)$ is indeed a p.m.f. Also $f(x)$ for any non-negative integer x is seen to be the $(x + 1)$st term in the expansion of the binomial $p^r (1 - q)^{-r}$ with a negative index Hence the name 'negative binomial'.

distribution on the basis of Table 3.9. ...

The values are recorded correct to one-tenth of a cm. Hence 144·6 represents any value between 144·55 and 144·65. Similarly, 149·5 represents any value between 149·45 and 149·55. Thus the class 144·6–149·5 really stands for the *class-interval* 144·55–149·55. Similar is the case for the other classes. 144·6 and 149·5 are called the lower and upper *class-limits* for the first class, while 144·55 and 149·55 are the corresponding *class-boundaries.* One should state the class-boundaries, rather than the class-limits, while drawing up the frequency distribution of a continuous variable. Table 3.10 shows the frequency distribution in terms of the absolute frequencies, relative frequencies and cumulative frequencies. It should be noted that the cumulative frequencies of the *less-than type* correspond to the *upper* class-boundaries ; for instance, the third one, 28, is the number of persons with height 159·55 cm. or less. Similarly, the cumulative frequencies of the *greater-than type* correspond to the *lower* class-boundaries.

TABLE 3.10

FREQUENCY DISTRIBUTION OF HEIGHT FOR 177 INDIAN ADULT MALES

Height (cm.) class-interval	Frequency	Relative frequency	Cumulative frequency 'Less-than'	Cumulative frequency 'Greater-than'
144·55—149·55	1	0·0057	1	177
149·55—154·55	3	0·0169	4	176
154·55—159·55	24	0·1356	28	173
159·55—164·55	58	0·3277	86	149
164·55—169·55	60	0·3390	146	91
169·55—174·55	27	0·1525	173	31
174·55—179·55	2	0·0113	175	4
179·55—184·55	2	0·0113	177	2
Total	177	1·0000	—	—

If the classes be of varying width, then the different class-frequencies will not be comparable. Comparable figures can be

Example 8.3 Fit a normal distribution to the frequency distribution of height of Indian adult males given in Table 3.10. Also draw the fitted curve over the histogram of the observed distribution.

For the distribution of height of Indian adult males, the mean and standard deviation were found to be

$$\bar{x} = 164 \cdot 734 \text{ cm. and } s = 5 \cdot 472 \text{ cm.}$$

Here $$n = 177 \text{ and } n/s = 32 \cdot 3462 .$$

TABLE 8.6

FITTING A NORMAL DISTRIBUTION TO THE HEIGHT-DISTRIBUTION OF INDIAN ADULT MALES (TABLE 3.10)

Height (cm.) x (1)	$\tau = (x - \bar{x})/s$ (2)	$\phi(\tau)$ (3)	Ordinate $= \frac{n}{s}\phi(\tau)$ (4)	$\Phi(\tau)$ (5)	$\Delta\Phi(\tau)$ (6)	Expected frequency $n \times \Delta\Phi(\tau)$ (7)	Observed frequency (8)
$-\infty$	$-\infty$	0	0	0			
					0·0001126*	0·020	0
144·55	–3·689	0·0004424	0·0143	0·0001126			
					0·0026478	0·469	1
149·55	–2·775	0·0084874	0·2745	0·0027604			
					0·0286123	5·064	3
154·55	–1·861	0·0706097	2·2839	0·0313727			
					0·1404492	24·860	24
159·55	–0·947	0·2547828	8·2412	0·1718219			
					0·3146168	55·687	58
164·55	–0·034	0·3987070	12·8966	0·4864387			
					0·3241316	57·371	60
169·55	0·880	0·2708640	8·7614	0·8105703			
					0·1530213	27·085	27
174·55	1·794	0·0798081	2·5815	0·9635916			
					0·0330236	5·845	2
179·55	2·708	0·0101984	0·3299	0·9966152			
					0·0032381	0·573	2
184·55	3·621	0·0005673	0·0183	0·9998533			
					0·0001467**	0·026	0
∞	∞	0	0	1			
Total	—	—	—	—	1·0000000	177·000	177

*It is the probability $P[x \leq 144{\cdot}55]$. **It is the probability $P[x \geq 184{\cdot}55]$.

With these, we can now compute the expected frequencies for the different class-intervals and the ordinates at the class-boundaries in the manner explained above. In the tables, $\phi(\tau)$ and $\Phi(\tau)$ are given for values of τ at intervals of 0·01 while in the present case we have taken $\tau = (x - \bar{x})/s$ correct to 3 decimal places. For obtaining $\phi(\tau)$ and $\Phi(\tau)$ for these values, we have used linear interpolation.

The agreement between the observed and the expected series of frequencies would seem to be fairly good. This agreement will also be apparent from Fig. 8.2, where we have the fitted normal curve, obtained on the basis of col. (4) of Table 8.6, superimposed on the histogram of the observed distribution.

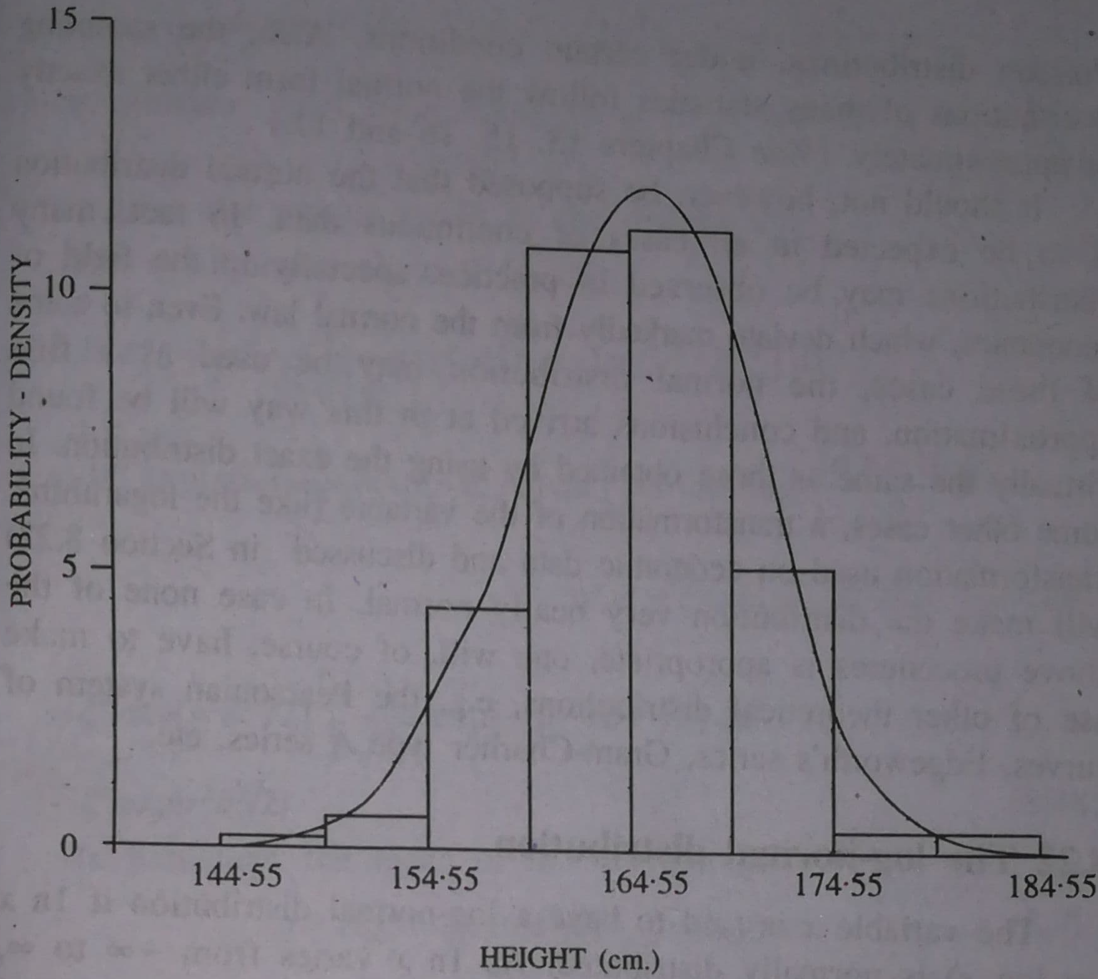

Fig. 8.2. Fitted normal curve together with the histogram of the height-distribution of Indian adult males (Table 3.10).

8.21 Importance of the normal distribution in statistics

The normal distribution plays a very important rôle in statistical theory and its applications. As we have already seen, it has some very simple properties which make it comparatively easy to deal with. Consequently, it will be a distinct advantage if in any case the population distribution of the variable under consideration may be assumed to be of the normal type. Generally, such an assumption is found legitimate in most cases of data arising from biological and psychological measurements. Under certain conditions, it can also be shown that the distribution of errors of observation in repeated measurements on a physical constant may be supposed to be normal. Such conditions being more or less valid in the field of manufacturing industry as well, most data arising there are also found to follow the normal law. Moreover, as we saw earlier, it serves as an approximation to the binomial and

Type VII. This occurs when $b_1 = 0$, and b_0 and b_2 are of the same sign. The equation to the curve is

$$f(x) = C\left(1+\frac{x^2}{a^2}\right)^{-m}, \quad -\infty < a < \infty, \qquad \ldots (8.98)$$

the origin being at the mean.

This is also symmetrical about the origin and is transformed into the beta form under the substitution $z = \left(1+\frac{x^2}{a^2}\right)^{-1}$. This is always bell-shaped.

The *normal curve* is also a transition type of the Pearsonian family and is obtained when $b_1 = b_2 = 0$.

It can be shown that b_0, b_1 and b_2, and hence κ, can be expressed in terms of β_1 and β_2. Thus the curves of the Pearsonian family can be specified by the β_1 and β_2 criteria.

Writing the differential equation in the form

$$\frac{df(x)}{dx} = \frac{x f(x)}{b_0 + b_1 x + b_2 x^2} \quad \text{(origin at mode, } \alpha\text{)},$$

we have

$$\frac{d^2 f(x)}{dx^2} = \frac{d}{dx}\left(\frac{xf(x)}{b_0 + b_1 x + b_2 x^2}\right) = \frac{xf(x)}{(b_0 + b_1 x + b_2 x^2)^2}\{(1-b_2)x^2 + b_0\}.$$

Thus each curve of the Pearsonian family has two points of inflection, given by

$$x = \pm\sqrt{\frac{b_0}{b_2 - 1}}, \qquad \ldots (8.99)$$

which are equidistant from the mode.

Questions and exercises

8.1 Explain the meaning and utility of theoretical distributions, and indicate the relevance of probability-mass and probability-density functions.

8.2 Derive the hypergeometric distribution from a suitable probability model. Also obtain its mean and s.d.

8.3 Derive the binomial distribution from a suitable probability model. Also indicate how this may be looked upon as a limiting form of the hypergeometric distribution. Obtain the mean and the s.d. of the distribution.

8.4 Derive the Poisson distribution from a suitable probability model and also as a limiting form of the binomial distribution. Give examples of data for which the Poisson distribution is expected to give a good fit.

8.5 Show that the normal distribution may be looked upon as a limiting form of the binomial and Poisson distributions. What are the important properties of this distribution ? Account for the importance of the normal distribution in statistical thoery and practice.

8.6 Determine the modes of the binomial and Poisson distributions. Show that the mode coincides with the mean when mp or λ (as the case may be) is an integer.

Partial ans. The modes are the highest integers contained in $(m+1)p$ and λ.

8.7 Let the intensity of accident-proneness, λ, of workmen follow a gamma distribution with p.d.f. $g(\lambda)=\dfrac{\gamma^{\alpha}}{\Gamma(\alpha)}\exp[-\gamma\lambda]\,\lambda^{\alpha-1}$, $0<\lambda<\infty$, and let the number of accidents made by a workman whose intensity of accident-proneness is λ follow a Poisson distribution with p.m.f. $f(x|\lambda)=\exp[-\lambda]\,\dfrac{\lambda^x}{x!}$, $x=0, 1, 2, \ldots\ldots$. Show that the number of accidents x, made by a workman of unknown accident-proneness, follows a negative binomial distribution.

8.8 Show that the cumulative probability of the binomial distribution may be expressed in the form

$$\sum_{x=0}^{k}\binom{m}{x}p^x q^{m-x}=\frac{1}{B(m-k,\,k+1)}\int_0^q z^{m-k-1}(1-z)^k\,dz$$

and that of the Poisson distribution in the form

$$\sum_{x=0}^{k}\exp[-\lambda]\frac{\lambda^x}{x!}=\frac{1}{\Gamma(k+1)}\int_{\lambda}^{\infty}\exp[-z]z^k\,dz.$$

8.9 Suppose a certain type of event occurs according to a Poisson process with mean rate λ per unit of time, so that the number of occurrences of the event in a time interval of length t is a Poisson random variable with mean λt. Show that the distribution of the waiting time till the first occurrence of the event is exponential with mean $\frac{1}{\lambda}$ and the distribution of the waiting time till the rth occurrence is gamma with parameters (λ, r).

8.10 (*a*) Obtain the moment-generating function of the Poisson distribution with parameter λ. Hence obtain the mean, μ_2, μ_3 and μ_4 for this distribution.

(*b*) Obtain the moment-generating function of the negative binomial distribution and hence determine its first four moments.

8.11 The *Pascal distribution* is defined by

$$f(x) = \frac{1}{1+\mu}\left(\frac{\mu}{1+\mu}\right)^x, \quad x = 0,1,2,\ldots\ldots,$$

where $\mu > 0$.

Find the mean and variance of the distribution.

8.12 Suppose 5% of the inhabitants of Calcutta are cricket fans. Determine approximately the probability that a sample of 100 inhabitants will contain at least 8 cricket fans? *Ans.* 0.126.

8.13 The probability of getting no misprint in a page of a book is 0.14. What is the probability that a page contains more than 2 misprints? (State the assumption you make.)

Ans. 0.31 (under proper assumption).

8.14 A poison distribution has a double mode at $x = 2$ and $x = 3$. What is the probability that x will have one or the other of the two values? *Ans.* 0.224.

8.15 Show that a random variable x distributed in the exponential form has the *lack of memory property* :

$$P\,[x > s + t \mid x > t] = P\,[x > s], \quad \text{for } s, t > 0,$$

This means that under the condition that an item survives to time t, the probability of surviving a further time s is the same as the probability of surviving to time s in the first place. It does not depend on t.

8.16 The continuous distribution with p.d.f.

$$f(x) = \frac{\alpha\, x_0^{\alpha}}{x^{\alpha+1}} \quad \text{if } x > x_0$$
$$= 0 \quad \text{otherwise,}$$

where $\alpha > 0$, is called a *Pareto distribution* and is found to be appropriate for variables like income or wealth per family in a community. Find the mean and variance of the distribution (assuming $\alpha > 2$) and also the distribution function.

8.17 Starting from an appropriate differential equation, obtain the curves of the Pearsonian system. Discuss their important properties.

8.18 Show that for a symmetrical probability distribution (either discrete or continuous), all odd-order central moments are equal to zero.

8.19 A continuous random variable x having values only between 0 and 4 has the density function $f(x) = \frac{1}{2} - ax$. Evaluate a.

8.20 Find the mean and variance of each of the following continuous probability distributions :

(i) $f(x) = \alpha \exp(-\alpha x), \; x \geq 0$ and $\alpha > 0$;

(ii) $f(x) = \frac{1}{2} \exp(-|x|), \; -\infty < x < \infty$.

8.21 Find the m.g.f. of the normal distribution with mean μ and variance σ^2. Hence show that

$$\mu_{2r+1} = 0$$

while
$$\mu_{2r} = (2r-1)\,\mu_{2r-2}\,\sigma^2.$$

Indicate how the m.g.f. of $N(\mu, \sigma^2)$ enables us to evaluate the moments of a lognormal distribution.

8.22 The life (in hours) of electronic tubes of a certain type is supposed to be normally distributed with $\mu = 155$ hr. and $\sigma = 19$ hr. What is the probability that the life of a tube will be

(a) between 136 hr. and 174 hr. ?
(b) between 117 hr. and 193 hr. ?
(c) less than 117 hr. ?
(d) more than 193 hr. ?

If a sample of 200 tubes is taken, how many are expected to be in each of the above groups ?

Partial ans. The probabilities are :
(a) 0·68 ; (b) 0·96 ; (c) 0·02 ; (d) 0·02.

8.23 The results of a particular examination are shown below in summary form :

Result	Percentage of candidates
Passed with distinction	15
Passed without distinction	42
Failed	43
Total	100

It is known that a candidate gets plucked if he obtains less than 40 marks (out of 100), while he must obtain at least 75 marks in order to pass with distinction. Hence determine the mean and s.d. of the distribution of marks, assuming it is of the normal type.

Ans. $\mu = 45{\cdot}09$; $\sigma = 28{\cdot}86$.

8.24 Show that the mean deviation about mean of a normal distribution is $\sqrt{\frac{2}{\pi}} \times \sigma$, σ being the s.d. of the distribution.

8.25 If $\log x$ is normally distributed with $\mu = 1$ and $\sigma^2 = 4$, find $P[\frac{1}{2} < x < 2]$.

Ans. 0·106.

8.26 There are 600 commerce students in the post-graduate classes of a university, and the probability for any student to need a copy of a particular text-book from the university library on any day is 0·05. How many copies of the book should be kept in the university library so that the probability may be greater than 0·90 that none of the students needing a copy from the library has to come back disappointed ? (Use the normal approximation to the binomial probability law.)

Ans. At least 37 copies.

8.27 Suppose the life-time (in hours) of a radio tube of a certain type obeys the exponential law $f(x) = \frac{1}{\lambda}\exp[-x/\lambda]$, $x > 0$, with $\lambda = 900$. A company producing tubes wishes to guarantee for the articles a certain life-time. For how many hours should the tube be guaranteed to function to achieve a probability of 0·90 that it will function for (at least) the number of hours guaranteed.

Ans. 95 hours.

8.28 For the continuous probability distribution

$$f(x) = \alpha \exp[-\alpha (x - \theta)], \quad \theta < x < \infty,$$

where $\alpha > 0$, find the moment-generating function. Obtain the mean, variance, β_1 and β_2 of the distribution.

8.29 In the course of an experiment, 15 mosquitoes were put in each of 120 jars and were next subjected to a dose of D.D.T. After 4 hours the number alive in each jar was counted and the following frequency distribution was obtained :

No. of mosquitoes alive	0	1	2	3	4	5	6	7	8
Frequency (no. of jars)	2	12	14	22	28	17	13	10	2

Find the frequencies that one would expect on the assumption that each mosquito has a common probability of survival.

8.30 When the first proof of a book containing 250 pages was read, the following distribution of misprints was obtained :

No. of misprints per page	Frequency
0	139
1	76
2	28
3	4
4	2
5	1
Total	250

Fit a Poisson distribution to the above data.

8.31 A telephone switch-board handles 720 calls on the average during a rush hour. The board can make 15 connections per minute. Estimate the probability that the board will be overtaxed during any minute in the rush hour. *Ans.* 0·156.

8.32 The following distribution relates to the number of accidents to 647 women working on H.E. (high explosive) shells during a 5-week period (given by Greenwood and Yule in *J.R.S.S.*, 1920). Show that a negative binomial distribution, rather than a Poisson distribution, gives a very good fit to the data. How would you explain this ?

Number of accidents	0	1	2	3	4	5
Frequency	447	132	42	21	3	2

[*Hint* : Refer to the result of *Exercise* 8.7.]

8.33 A car hire firm has two cars, which are hired out by the day. It has been found that the number of demands for cars of the firm on any day has a Poisson distribution with mean 1·5.

(*a*) Calculate the proportion of days on which neither car is used and the proportion of days on which some demand is refused.

(*b*) If the two cars are used an equal number of times on the average, on what proportion of days is a given one of the cars not in use ?

(*c*) How many cars should the firm have so as to meet all demands on approximately 98% of days ?

Ans. (*a*) 0·223, 0·191 ; (*b*) 0·390 ; (*c*) 4.

8.34 The following is the frequency distribution of right-hand grip for 345 European males :

Right-hand grip (in lb.)	Frequency
29·5— 39·5	1
39·5— 49·5	2
49·5— 59·5	12
59·5— 69·5	52
69·5— 79·5	99
79·5— 89·5	101
89·5— 99·5	55
99·5— 109·5	17
109·5— 119·5	5
119·5— 129·5	1
Total	345

Find the expected frequencies for the above classes assuming that the population distribution of right-hand grip is normal. Draw the fitted curve and the histogram of the observed distribution on the same graph paper.

Suggested Reading

1. Bhattacharyya, G.K. and Johnson, R.A. *Statistical Concepts and Methods* (Chs. 5, 7). John Wiley, 1977.
2. Elderton, W.P. and Johnson, N.L. *Systems of Frequency Curves.* Cambridge University Press, 1969.
3. Feller, W. *An Introduction to Probability Theory and its Applications, vol. I* (Chs. 6, 7). John Wiley, 1968, and Wiley Eastern, 1972.
4. Freund, J.E. *Mathematical Statistics* (Chs. 5, 6). Prentice-Hall and Prentice-Hall of India, 1992.
5. Hoel, P.G., Port, S.C. and Stone, C.J. *Introduction to Probability Theory* (Chs. 3-5, 7). Houghton Mifflin, 1971, and Universal Book Stall.
6. Mukhopadhyay, P. *Mathematical Statistics* (Chs. 5. 6). New Central Book Agency, 1996.
7. Parzen, E. *Modern Probability Theory and its Applications.* (Chs. 3-6). John Wiley, 1960, and Wiley Eastern, 1972.
8. Pitman, J. *Probability* (Chs. 2, 3). Springer-Verlag, 1993, and Narosa, 1996.
9. Rohatgi, V.K. *An Introduction to Probability Theory and Mathematical Statistics* (Chs. 2, 3, 5). John Wiley, 1976, and Wiley Eastern, 1985.