The Poisson Distribution and Symbolic Computations

Please attach a cover page (-10%).

Problem 1 (75% = 3 x 25%): Exercises 6.1 and 6.2 from Bulmer, Dover (1979), p. 104, and **Exercise 6.8** from Bulmer, Dover (1979), p. 105**.**

Problem 2 (25%):
Calculate the mean and variance of the observations in Table 14 in Bulmer, Dover (1979), p. 95. Comment.

```
(* BIOEN 3070/6070: Introduction to Statistics for Bioengineers *)
```



```
(* The Poisson Distribution and Symbolic Computations *)
```

```
(* Problem 3 (100% Extra Credit):
 Use symbolic computations to show that the
  mean of a given set of observations minimizes the variance. *)
```

```
? D
```

D[f, x] gives the partial derivative $\partial f / \partial x$.
D[f, {x, n}] gives the multiple derivative $\partial^n f / \partial x^n$.
D[f, x, y, …] differentiates f successively with respect to x, y, ….
D[f, {{x_1, x_2, …}}] for a scalar f gives the vector derivative $(\partial f / \partial x_1, \partial f / \partial x_2, \ldots)$.
D[f, {*array*}] gives a tensor derivative. ≫

```
? Solve
```

Solve[*expr*, *vars*] attempts to solve the system *expr* of equations or inequalities for the variables *vars*.
Solve[*expr*, *vars*, *dom*] solves over the domain *dom*. Common choices of *dom* are Reals, Integers, and Complexes. ≫

Exercises

6.1. In a table of random numbers, digits are arranged in groups of four. One proposed test of randomness is the *poker* test in which the groups are classified as

(*i*) all four digits the same
(*ii*) three digits the same, and one different
(*iii*) two of one digit and two of another
(*iv*) two of one digit and two different digits
(*v*) all four different

and the observed frequencies compared with expectation. Find the probabilities of each of these events. (P. P. P., Hilary, 1963.)

6.2. When Mendel crossed a tall with a dwarf strain of pea and allowed the hybrids to self-fertilise he found that $\frac{3}{4}$ of the offspring were tall and $\frac{1}{4}$ dwarf. If 9 such offspring were observed, what is the chance (*a*) that less than half of them would be tall?; (*b*) that all of them would be tall?

6.3. Find the mean and variance of the observed frequency distribution and of the theoretical probability distribution in Table 11 on p. 82 directly, using the exact probabilities in fractional form for the latter; compare these values with those obtained from the formula for the binomial distribution with $n = 5$ and $P = \frac{1}{2}$.

6.4. Weldon threw 12 dice 4096 times, a throw of 4, 5 or 6 being called a success, and obtained the following results:

No. of successes	0	1	2	3	4	5	6	7	8	9	10	11	12	Total
Frequency	0	7	60	198	430	731	948	847	536	257	71	11	0	4096

(*a*) Calculate the mean of the distribution and compare it with that of a binomial distribution with $P = \frac{1}{2}$; find the observed proportion of successes per throw (p).

(*b*) Calculate the variance of the distribution and compare it with that of a binomial distribution (*i*) with $P = \frac{1}{2}$; (*ii*) with $P = p$.

(*c*) Fit a binomial distribution, (*i*) with $P = \frac{1}{2}$, (*ii*) with $P = p$.

6.5. Evaluate the variance of the sex ratios in Table 1 on p. 3 (*a*) in the Regions of England, (*b*) in the rural districts of Dorset, and compare them with the theoretical variances of a proportion (*a*) with $n = 100{,}000$, (*b*) with $n = 200$.

6.6. The following is one of the distributions of yeast cells observed by 'Student'. Fit a Poisson distribution to it.

No. of yeast cells	0	1	2	3	4	5	6	Total
No. of squares with this no. of cells in it	103	143	98	42	8	4	2	400

6.7. The frequency of twins in European populations is about 12 in every thousand maternities. What is the probability that there will be no twins in 200 births, (*a*) using the binomial distribution, (*b*) using the Poisson approximation?

6.8. When a bacterial culture is attacked by a bacterial virus almost all the bacteria are killed but a few resistant bacteria survive. This resistance is hereditary, but it might be due *either* to random, spontaneous mutations occurring before the viral attack, *or* to mutations directly induced by this attack. In the latter case the number of resistant bacteria in different but similar cultures should follow the Poisson distribution, but in the former case the variance should be larger than the mean since the number of resistant bacteria will depend markedly on whether a mutation occurs early or late in the life of the culture. In a classical experiment Luria & Delbruck (1943) found the following numbers of resistant bacteria in ten samples from the same culture: 14, 15, 13, 21, 15, 14, 26, 16, 20, 13, while they found the following numbers in ten samples from different cultures grown in similar conditions: 6, 5, 10, 8, 24, 13, 165, 15, 6, 10. Compare the mean with the variance in the two sets of observations. Comment.

6.9. Calculate the mean and the standard deviation of the distribution in Table 16 on p. 99 and show that they are nearly equal.

6.10. If traffic passes a certain point at random at a rate of 2 vehicles a minute, what is the chance that an observer will see no vehicles in a period of 45 seconds?

Problems

6.1. The rth factorial moment, $\mu_{[r]}$, is the Expected value of $X(X-1)\ldots(X-r+1)$. Find the rth factorial moment of the binomial distribution (*a*) by direct calculation, (*b*) from the p.g.f. and hence find its mean, variance, skewness and kurtosis. (Cf. Problem 5.9.)

6.2. Do the same calculations for the Poisson distribution.

6.3. If X and Y are independent Poisson variates with means μ and ν, show that

$$\text{Prob}\,[X+Y=c] = \sum_{i=0}^{c} \text{Prob}\,[X=i \;\&\; Y=c-i] = \frac{e^{-(\mu+\nu)}(\mu+\nu)^c}{c!}.$$

6.4. Suppose that a suspension of cells contains μ cells per c.c. so that the number of cells in 1 c.c. follows a Poisson distribution with mean μ; and suppose that the probability that a cell will survive and multiply to form a visible colony is P, so that the number of colonies formed from a fixed number of cells follows a binomial distribution. Show that the number of colonies formed from 1 c.c. of the suspension follows a Poisson distribution with mean μP.