Computational Statistics and Data Analysis (MVComp2)

Exercise 3

Lecturer Tristan Bereau

Semester Wi23/24
Due Nov. 9, 2023, 23:59

1 Polluted water (2 points)

A particular concentration of a chemical found in polluted water has been found to be lethal to 20% of the fish that are exposed to the concentration for 24 hours. Twenty fish are placed in a tank containing this concentration of chemical in water.

(a) Find the probability that at least 14 survive.

(b) Find the mean and variance of the number that survive.

2 Distribution means (4 points)

(a) If X is a random variable with a geometric distribution $P(X = x) = (1-p)^{k-1}p$, prove that the mean is given by

$$E[X] = \frac{1}{p}$$

(b) If X is a random variable with a Poisson distribution $P(X = x) = \frac{\lambda^x}{x!}\mathrm{e}^{-\lambda}$, prove that the mean is given by

$$E[X] = \lambda$$

3 Conjugate priors (2 points)

Consider the Beta distribution with parameters α and β

$$\mathrm{Beta}(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}$$

and the Binomial distribution for N observations

$$P(X = k) = \binom{N}{k}\theta^k(1-\theta)^{N-k}.$$

Show that the Beta distribution is a conjugate prior to the Binomial distribution. What are the parameters of the resulting distribution?

4 The de Moivre–Laplace theorem (2 points)

The de Moivre–Laplace theorem states that the normal distribution may be used as an approximation to the Binomial distribution under certain conditions.[1] Let's find out!

Consider a random variable, X. The probability of getting k successes in n independent Bernoulli trials with probability p is given by the Binomial

$$P(X = k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}.$$

The thoerem states that the Binomial probability will converge to the normal probability density function, $\mathcal{N}(np, \sqrt{np(1-p)})$, as n becomes large and for p away from 0 or 1.

We will verify that the standardized random variable, $\frac{X-np}{\sqrt{np(1-p)}}$, approaches the standard normal, $\mathcal{N}(0,1)$, as n grows larger. Consider $p = 0.5$ and the cases $n = \{2, 5, 10, 50, 100\}$. Plot all curves on the same graph to show that they overlap.

[1] This is a special case of the central limit thoerem. A famous physical realization of the de Moivre–Laplace theorem is the Galton box, also called bean machine.