CSE Exercises - Week 4

(1) In this exercise, we show analytically and computationally that adding random variables is different from adding distributions and so the two operations should not be confused.

Consider two normal distributions, Normal(μ_1, σ_1^2) and Normal(μ_2, σ_2^2).

(a) Adding random variables.

Let X and Y be independent random variables such that $X \sim$ Normal(μ_1, σ_1^2) and $Y \sim$ Normal(μ_2, σ_2^2). Let $Z = X + Y$.

(i) Show that $E(Z) = \mu_1 + \mu_2$, and $V(Z) = \sigma_1^2 + \sigma_2^2$.

(ii) Now let $\mu_1 = -2$, $\mu_2 = 2$ and $\sigma_1^2 = \sigma_2^2 = 1$. We will use simulation to show, as expected, that the distribution of Z is Normal$(0, 2)$:

• Let $n = 10,000$.

• Generate $X_1, \ldots, X_n \sim$ Normal$(-2, 1)$.

• Generate $Y_1, \ldots, Y_n \sim$ Normal$(2, 1)$.

- Compute $Z_i = X_i + Y_i$, $i = 1, \ldots, n$.

- Plot a density histogram for $Z_1, \ldots, Z_n$.

- On the same plot, superimpose the curve for the Normal$(0, 2)$ density.

(b) Adding distributions.

Let $f(x; \mu_1, \sigma_1^2)$ denote the Normal(μ_1, σ_1^2) density and $f(x; \mu_2, \sigma_2^2)$ denote the Normal(μ_2, σ_2^2) density. Define a 2-component normal mixture density as

$$g(x) := \omega f(x; \mu_1, \sigma_1^2) + (1-\omega) f(x; \mu_2, \sigma_2^2)$$

where $\omega \in (0, 1)$.

(i) Show that $g(x)$ integrates to 1 (as it should to be a proper density).

(ii) Let X be a random variable with density $g(x)$. Show that
$$E(X) = \omega \mu_1 + (1-\omega) \mu_2$$
and
$$V(X) = \omega \sigma_1^2 + (1-\omega) \sigma_2^2 + \omega(1-\omega)(\mu_1 - \mu_2)^2.$$
Notice the differences compared to part (a)(i).

(iii) We can generate a random variable, X, with density $g(x)$ by the following algorithm:

Generate $U \sim$ Uniform $(0,1)$,
If $U \leq \omega$
     generate $X \sim$ Normal (μ_1, σ_1^2),
else
     generate $X \sim$ Normal (μ_2, σ_2^2).
Return X.

Now let $\mu_1 = -2$, $\mu_2 = 2$ and $\sigma_1^2 = \sigma_2^2 = 1$ as in part (a)(ii). We will use simulation to show that, unlike part(a)(ii), $g(x)$ is not a normal density.

- Let $n = 10,000$.

- Implement the algorithm given above and use it to generate $X_1, \ldots, X_n$ with density $g(x)$.

- Plot a density histogram for $X_1, \ldots, X_n$.

- On the same plot, superimpose the curve for $g(x)$.

②   Page 142 Exercise 105.

③ In this exercise, we investigate the approximation of the Binomial(n, θ) distribution by the Poisson(λ) distribution with $\lambda = n\theta$.

(a) Let $n = 20$ and $\theta = 0.1$. Produce a plot of the Binomial$(20, 0.1)$ PMF, similar to the one given in Figure 6.13. On the same plot, superimpose the Poisson(λ) PMF (use a different symbol or colour so that the two PMFs are easily distinguishable).

(b) Repeat part (a) with $\theta = 0.5$.

(c) Repeat part (a) with $\theta = 0.9$.

(d) What can you conclude from the plots from (a), (b) and (c)?

(e) Describe how, for $\theta > 0.5$, the approximation of the binomial distribution by the Poisson distribution can be made as accurate as for $\theta < 0.5$.

④ In this exercise, we investigate the approximation of the Binomial(n, θ) distribution by the Normal(μ, σ^2) distribution with $\mu = n\theta$ and $\sigma^2 = n\theta(1 - \theta)$, when n is large and p is not too close to 0 or 1.

It may come as a surprise that a discrete distribution can be approximated by a continuous one but this exercise shows that this is possible.

(a) Let $n = 20$ and $\theta = 0.1$. Produce a plot of the Binomial $(20, 0.1)$ PMF, similar to the one given in Figure 6.13. Compute the Normal $(2, 1.8)$ densities for $x = 0, 1, \ldots, 20$. Superimpose these onto the same plot.

(b) Repeat part (a) with $\theta = 0.5$.

(c) Repeat part (a) with $\theta = 0.9$.

(d) What can you conclude from the plots from (a), (b) and (c)?

(5) The Poisson (λ) distribution can also be approximated by the Normal (μ, σ^2) distribution with $\mu = \lambda$ and $\sigma^2 = \lambda$, when λ is large. To see this,

(a) Let $\lambda = 2$ and plot the Poisson (2) PMF for a range of x values, say $x = 0, 1, \ldots, 10$. Compute the Normal $(2, 2)$ densities for the same x values and superimpose them onto the plot.

(b) Repeat part (a) with $\lambda = 20$ and for $x = 0, 1, \ldots, 40$.

(c) Repeat part (a) with $\lambda = 100$ and for $x = 70, 71, \ldots, 130$.

(d) What can you conclude from the plots?

(6) Like the normal random varibles, Poisson random variables have the property that a sum of independent Poisson random variables is also a Poisson random variable. Thus, if $X_1, ..., X_n$ are independent random variables such that $X_i \sim$ Poisson (λ_i), then

$$X_1 + \cdots + X_n \sim \text{Poisson } (\lambda_1 + \cdots + \lambda_n).$$

We shall demonstrate this for a sum of two independent Poisson random variables.

- Let $n = 10,000$.

- Generate $X_1, ..., X_n \sim$ Poisson (3).

- Generate $Y_1, ..., Y_n \sim$ Poisson (6).

- Compute $Z_i = X_i + Y_i$, $i = 1, ..., n$.

- Compare the relative frequencies from the resulting $Z_1, ..., Z_n$ against the PMF of the Poisson (9) distribution.

- Estimate $E(Z)$ and $V(Z)$ using $Z_1, ..., Z_n$ and comment on your estimated values.