Week 2: Probability review
Bernoulli, binomial, Poisson, and normal distributions

In this exercise, we study Bernoulli, binomial, Poisson, and normal random variables (RVs) as well as the relations between these probability distributions. Let us start by reviewing each of them.

A *Bernoulli* RV L models experiments where success happens with probability p and failure with probability $1-p$, for instance, a coin toss. Success is indicated by $L = 1$ and failure by $L = 0$. These "experiments" are sometimes referred to as "Bernoulli experiments" or "Bernoulli trials". Therefore, the *probability mass function* (pmf) of L is

$$\mathbb{P}[L = 0] = 1 - p, \quad \mathbb{P}[L = 1] = p, \tag{1}$$

and $\mathbb{P}[L = k] = 0$ for all other integers k.

A *binomial* RV with parameters (n, p) counts the number of successes in n independent Bernoulli trials that succeed with probability p. Hence, we can write a binomial RV B as

$$B = \sum_{i=1}^{n} L_i, \tag{2}$$

where the $\{L_i\}$ are independent Bernoulli RVs with pmfs as in (1). Clearly, B can only takes integer values between 0 and n. The pmf of a binomial RV is easily derived by noting that for B to be equal an integer $k \in [0, n]$, we must have exactly k successful Bernoulli trials—something that happens with probability p^k—and exactly $n - k$ failed trials—which happens with probability $(1 - p)^{n-k}$. Moreover, we need to consider that there are $\binom{n}{k}$ different ways in which this could happen. Thus, the binomial pmf is

$$\mathbb{P}[B = k] = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}, \quad \text{for } k = 0, 1, \ldots, n, \tag{3}$$

and $\mathbb{P}[B = k] = 0$ otherwise.

A *Poisson* RV P takes values in the nonnegative integers. As we will see in part C, Poisson RV are used to describe/model "counting processes", similar to how the binomial counts the number of successes in Bernoulli trials. We say that P is Poisson with parameter λ if its pmf is

$$\mathbb{P}[P = k] = e^{-\lambda} \frac{\lambda^k}{k!}, \quad \text{for } k = 0, 1, \ldots, \tag{4}$$

and $\mathbb{P}[P = k] = 0$ otherwise.

In contrast to the previous RVs, a *normal* or *Gaussian* RV X can take on any real value (instead of just 0 or 1 for the Bernoulli, integers between 0 and n for the binomial, or nonnegative integers for the Poisson). We therefore say X is a *continuous* RV. Since we can no longer specify probabilities for each single point (as we have done until now), probabilities are now described using a *probability* *density function* (pdf). The pdf of a normal RV with mean μ and variance σ^2 is

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \text{for } x \in \mathbb{R}. \tag{5}$$

Another important concept, especially for continuous RVs, is that of the *cumulative distribution function* (cdf), defined as the probability of an RV Y not exceeding y, i.e., $F_Y(y) = \mathbb{P}[Y \leq y]$. For non-negative discrete pdfs such as the Bernoulli, binomial and Poisson, we can write

$$F_Y(y) = \sum_{u=0}^{y} \mathbb{P}[Y = u]. \tag{6}$$

When the RV is continuous, as is the case of the normal/Gaussian RV, the sum in (6) is replaced by an integral as in

$$F_Y(y) = \int_{-\infty}^{y} f_Y(u)\, du. \tag{7}$$

Unlikely as it may seem, all these RV are closely related. We already saw the relation between the Bernoulli and the binomial and you will explore the other connections below.

A The Binomial RV. Prove that the expected value of a binomial RV B_n with parameters (n, p) is $\mathbb{E}[B_n] = np$ and that its variance is $\mathbb{E}\left[(B_n - \mathbb{E}\, B_n)^2\right] = np(1-p)$. Fix the expected value $\mathbb{E}[B_n] = np = 5$ and plot the pmf and cdf for $n = 6, 10, 20, 50$. Modify the value of p appropriately.

B Binomial and Poisson distributions. Prove that the expected value of a Poisson RV P with parameter λ is $\mathbb{E}[P] = \lambda$. Plot the pmf of a Poisson distribution with parameter $\lambda = 5$. Notice that this pmf is quite similar to the binomial pmf of Part A for large n. In fact, we can quantify this proximity by evaluating the following mean squared error (MSE):

$$\Delta(B_n, P) = \sum_{k=0}^{\infty} \left(\mathbb{P}[B_n = k] - \mathbb{P}[P = k]\right)^2 \mathbb{P}[P = k]. \tag{8}$$

Compute $\Delta(B_n, P)$ for $n = 6, 10, 20, 50$. Evaluating the MSE in (8) numerically requires truncating the infinite sum. You can neglect values k for which $\mathbb{P}[P = k] \leq 5 \times 10^{-2}$.

C Binomial and Poisson distributions again. Having noticed this interesting connection between the binomial and the Poisson distribution, consider binomial RVs B_n with parameters n and $p = \lambda/n$. Prove that as $n \to \infty$, the pmf of B_n converges to the pmf of P.

D Binomial and normal distributions. An important result in probability theory is the *central limit theorem* (CLT). The CLT concerns sums of independent identically distributed (i.i.d.) RVs Y_i with mean $\mathbb{E}[Y_i] = \mu$ and variance $\text{var}[Y_i] = \sigma^2$. Specifically, define

$$Z_n = \frac{\sum_{i=1}^{n} Y_i - n\mu}{\sigma\sqrt{n}}. \tag{9}$$

The CLT states that the pdf of Z_n is approximately normal (with zero mean and unit variance) for sufficiently large n. Formally,

$$\lim_{n\to\infty} \mathbb{P}\left[Z_n \le z\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2}\, du. \tag{10}$$

Recall that the binomial RV is a sum of i.i.d. Bernoulli RVs. Hence, it is possible to approximate the binomial cdf with a normal cdf. Do so for $p = 0.5$ and $n = 10, 20, 50$. Show the equations you used for the approximations and corresponding plots.

E Normal and Poisson approximations. In parts B and C, you showed that it is possible to approximate a binomial RV with a Poisson RV when n is large. But in part D, you showed that it is also possible to approximate a binomial RV with a normal RV. Clearly, Poisson and normal RVs are very different, but these results cannot contradict each other since you saw both are true. Explain why these two approximations are consistent with each other. The answer is *not* that the Poisson and normal are similar (they are not! E.g., the Poisson is a discrete distribution and the normal is a continuous distribution).