

Using Poisson random variables as approximations to Binomial random variables. These approximations work well if  $n$  is very large, and if  $npq$  is relatively close to 1. This might not be so satisfying for you, because there is not a hard-and-fast way of stating when the approximation works well. It just works better and better as  $n$  gets larger and larger, and also as  $npq$  is closer and closer to 1. Usually having  $npq$  within (say) a factor of 10 away from 1 is OK, e.g., between 0.1 and 10. I.e., a factor of 10 larger or smaller.

General idea: If  $n$  is large and  $npq$  is near 1, and if  $X$  is a Binomial random variable with parameters  $n$  and  $p$  (and here  $q = 1 - p$ ), and if  $Y$  is a Poisson random variable with  $\lambda = np$ , then  $X$  and  $Y$  have similar distributions for values (say) near the mean. The plots of the masses of each look like bell curves, so this intuitively makes sense. You can also write the mass of  $X$  and show it is pretty close to the mass of  $Y$ . In particular,  $E(X) = np = \lambda = E(Y)$ . So the mean values are exactly the same.

Let's try an example: Suppose that 1/10000 people own a purple car. Randomly sample 50000 people, with independent samples (e.g., no resampling, no sampling two people in the same family, etc.). Let  $X$  denote the number of people in the sample of 50000 who own a purple car. Then  $X$  has a Binomial( $n = 50000, p = 1/10000$ ) distribution. In particular,  $np = 5$ , i.e.,  $E(X) = 5$ . Now define  $Y$  to be a Poisson( $\lambda = 5$ ) random variable. Write the exact value of the probability that  $X = 7$ , and then use a Poisson approximation to get a good estimate for this value. Exact calculation:

$$P(X = 7) = p_X(7) = \binom{50000}{7} (1/10000)^7 (9999/10000)^{49993} = \binom{n}{7} p^7 q^{50000-7}.$$

This is very hard to compute on the calculator. E.g.,  $\binom{50000}{7} = (50000)(49999) \cdots (49994)/7!$ . If we do the calculation exactly, e.g., on a computer, we get

$$P(X = 7) = 0.104447996 \dots$$

We could instead use Poisson approximation:

$$P(Y = 7) = \frac{e^{-5} 5^7}{7!} = 0.104444863 \dots$$

The two values are very close, but the second value can be calculated much more readily on a handheld calculator.

We will learn other approximations later in the semester, as well, e.g., both the masses of Binomial and Poisson random variable can be calculated in many cases (approximately) using Normal random variables, which we have yet to learn about.