

Hotel bookings

Let X be the number of “no shows” that night. The natural distribution to hypothesize for this number is the Binomial distribution, and based on the description of the problem, we have that $X \sim \text{Bin}(12, .2)$. This immediately gives us the mean number of no shows,

$$\mu = E(X) = np = (12)(.2) = 2.4,$$

and the standard deviation of the number of no shows,

$$\sigma = SD(X) = \sqrt{np(1-p)} = \sqrt{(12)(.2)(.8)} = 1.39.$$

The hotel ends up with more customers than they can handle if 11 or 12 people show up, which corresponds to the number of no shows being 0 or 1. That is, we are interested in

$$\begin{aligned} P(\text{Too many customers}) &= P(X = 0 \text{ or } X = 1) \\ &= P(X = 0) + P(X = 1) \\ &= \binom{12}{0} (.2)^0 (.8)^{12} + \binom{12}{1} (.2)^1 (.8)^{11} \\ &= .069 + .206 = .275. \end{aligned}$$

We are making the usual Binomial process assumptions, which in this case correspond to the probability of a no show being the same for all reservations, and one reservation being a no show saying nothing at all about another being a no show. It is likely that both of these are violated: people who are coming from a longer distance probably have a higher probability of being a no show than people coming from shorter distances (violating the first assumption), people often travel together, and outside influences like weather can affect more than one reservation (violating the second assumption). Both of these violations lead to what is called *overdispersion*, where the variability in the number of successes is larger than would be implied by the Binomial distribution. This requires a more general family, and the *Beta-Binomial* family (of which the Binomial is a special case) is often used for this purpose.