

1. Prove that if  $A$  and  $B$  are independent, so are
  - (a)  $A$  and  $B^c$ ,
  - (b)  $A^c$  and  $B^c$ .
2. With election season nearing again, let's look (several years back) at the *fantastically* close race in 2000 in Florida between Al Gore and George W. Bush. Among the six million votes cast in Florida, it appears that the margin of victory may have been within just 500 votes.<sup>1</sup> To appreciate how extremely close this is, consider the following:
  - (a) Suppose a fair coin with Gore on one side and Bush on the other is tossed six million times, and each candidate gets a vote for each coin toss they win. At the end of the six million tosses, whoever has the most votes wins. What is the probability that, after tossing the coin six million times, the difference between the number of votes for Gore and Bush is less than or equal to 500?
  - (b) Repeat part a) if the coin is not fair, but instead lands Bush with probability .501 and Gore with probability .499. Are you surprised by the answer? What is going on here?
  - (c) Suppose a coin lands Bush with probability  $.5 + \epsilon$  and Gore with probability  $.5 - \epsilon$ , where  $\epsilon$  is some very small positive number. Find the value of  $\epsilon$  that gives Gore a 25 percent chance of winning. Sketch a picture of the distribution and the regions you are considering.

**A Sneak Peek:** When  $n$  is large enough such that  $10 \leq np \leq n - 10$ , we can use the normal distribution to approximate the binomial distribution. The normal density curve that best approximates the  $\text{Binomial}(n, p)$  probability distribution is the normal curve with mean  $\mu = np$  and standard deviation  $\sigma = \sqrt{np(1-p)}$ . Then, with a small *continuity correction*, we can approximate:

$$P(a \leq X \leq b) \approx P(a - \frac{1}{2} \leq N(\mu = np, \sigma = \sqrt{np(1-p)}) \leq b + \frac{1}{2}).$$

---

<sup>1</sup>Credit goes to Brian D. Jones, Kenyon College, for this problem.

**Computing Probabilities:** Open up Excel or LibreOffice<sup>2</sup>

Suppose we have a binomial distribution with  $n$  trials and the probability of success equal to  $p$ . We want to calculate some probabilities.

The syntax `=BINOMDIS(k, n, p, 0)` will calculate the probability that the number of success is exactly  $k$  (i.e. it will give  $b(n, p, k)$  using the notation from the textbook).

The syntax `=BINOMDIS(k, n, p, 1)` will calculate the probability that the number of successes is *less than or equal to*  $k$ . That is, for the binomial distribution with  $n$  trials and probability of success  $p$ , this expression calculates  $P(X \leq k)$ .

Suppose we want to ask: what is the value  $k$  for which  $P(X \leq k) = a$ , where  $a$  is some probability. For instance,  $P(X \leq k) = .85$  is asking: what is the number of successes,  $k$ , needed in order to ensure that 85 percent of the time we will have fewer than  $k$  successes?

The piece of code needed to compute this value  $k$  is: `=BINOM.INV(n,p,a)`.

---

<sup>2</sup>LibreOffice uses semi-colons in place of the commas. Eg. Your expressions look like: `=BINOMDIS(k; n; p; 0)`.