

## 1 Ball in Bins

You are throwing  $k$  balls into  $n$  bins. Let  $X_i$  be the number of balls thrown into bin  $i$ .

1. What is  $\mathbb{E}[X_i]$ ?
2. Compute  $\mathbb{E}[X_i^2]$ .
3. What is the expected number of empty locations?
4. What is the expected number of collisions?

## 2 How Many Queens?

You shuffle a standard 52-card deck, before drawing the first three cards from the top of the pile. Let  $X$  denote the number of queens you draw.

- (a) What is  $\mathbb{P}(X = 0)$ ?
- (b) What is  $\mathbb{P}(X = 1)$ ?
- (c) What is  $\mathbb{P}(X = 2)$ ?
- (d) What is  $\mathbb{P}(X = 3)$ ?
- (e) Do the answers you computed in parts (a) through (d) add up to 1, as expected?
- (f) Compute  $\mathbb{E}(X)$  from the definition of expectation.
- (g) Suppose we define indicators  $X_i$ ,  $1 \leq i \leq 3$ , where  $X_i$  is the indicator variable that equals 1 if the  $i$ th card is a queen and 0 otherwise. Compute  $\mathbb{E}(X)$ .

(h) Are the  $X_i$  indicators independent? Does this affect your solution to part (g)?

### 3 More Family Planning

(a) Suppose we have a random variable  $N \sim \text{Geom}(1/3)$  representing the number of children of a randomly chosen family. Assume that within the family, children are equally likely to be boys and girls. Let  $B$  be the number of boys and  $G$  the number of girls in the family. What is the joint probability distribution of  $B, G$ ?

(b) Given that we know there are 0 girls in the family, what is the most likely number of boys in the family?

(c) Now let  $X$  and  $Y$  be independent random variables representing the number of children in two independently, randomly chosen families. Suppose  $X \sim \text{Geom}(p)$  and  $Y \sim \text{Geom}(q)$ . Using their joint distribution, find the probability that the number of children in the first family ( $X$ ) is less than the number of children in the second family ( $Y$ ). (You may use the convergence formula for a Geometric Series:  $\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$  for  $|r| < 1$ )

(d) Show how you could obtain your answer from the previous part using an interpretation of the geometric distribution.