Sta-230: Probability **Spring 2017**

Key ideas for topic 4: February 2, 7

Hypergeometric distribution: The hypergeometric can be considered as sampling without replacement. Consider a bin with N balls of which M are red, the remaining balls are blue. You draw n balls and record the number that are red as m. This procedure is sampling without replacement and the distribution of m is hypergeometric

$$\mathbf{P}(m; N, M, n) = \frac{\binom{n}{m}\binom{N-M}{n-k}}{\binom{N}{n}}.$$

Multinomial distribution: The multinomial distribution is a generalization of the binomial distribution. Consider K possible categories (instead of two) with the probability of drawing an observation for each category as p_k where $k = 1, ..., K$ with $p_k \geq 0$ and $\sum_{k=1}^{K} p_k = 1$. The probability of drawing $\{x_1, x_2, ..., x_{K-1}, x_k\}$ observations from each category assuming we draw n observations is

$$\mathbf{P}(\{x_1, ..., x_K\}; n, \{p_1, ..., p_k\}) = \frac{n!}{\prod_k x_k!} \prod_k p_k^{x_k}.$$

Geometric distribution: Consider two possible events $X = \{0, 1\}$ (say heads and tails) that are Bernoulli with parameter p and you want to write down the distribution of the number of times T you need to draw an observation until you observe a 1. This distribution is the geometric distribution

$$\mathbf{P}(t; p) = (1-p)^t p,$$

this is the probability of observing t zeros and then a one.

Infinite outcome spaces: The geometric and Poisson distributions are examples of discrete distributions with infinite outcome spaces, the number of possible outcomes is infinity. A property of both spaces is that there will be zero probability outcomes

$$\lim_{t\to\infty} \left[\mathbf{P}(t; p > 0) = (1-p)^t p\right] = 0, \quad \lim_{k\to\infty} \left[\mathbf{P}(k; \lambda < \infty) = \frac{e^{-\lambda}\lambda^k}{k!}\right] = 0.$$

To see why the second limit is zero note the following approximation

$$n! \approx C \left(\frac{n}{e}\right)^n \sqrt{n}$$

so

$$\frac{\lambda^k}{k!} \approx C^{-1} \left(\frac{e\lambda}{k}\right)^k, \quad \lim_{k\to\infty} \left[\left(\frac{e\lambda}{k}\right)^k\right] = 0.$$