Lecture #23: Classical Probability Distributions
Key Concepts

Note: This material is ***for interest only*** — it will not be included in any assignment or test in this course.

Geometric Distributions

One Version

Situations Modelled: You repeatedly try to accomplish something. Each time you try, you succeed with probability p where $0 < p < 1$ — so that you fail, each time, with probability $1 - p$. You try up to K times — giving up, and stopping, if you failed K times.

What This is Called: The ***geometric distribution, truncated at K, with parameter p.***

Another Version

Situations Modelled: You repeatedly try to accomplish something. Each time you try, you succeed with probability p where $0 < p < 1$ — so that you fail, each time, with probability $1 - p$. You never stop trying before you succeed.

What This is Called: The ***geometric distribution with parameter p.***

Binomial Distributions

Situations Modelled: You are repeatedly trying to accomplish something — once again, succeeding each time with probability p, where $0 < p < 1$. This time, you make exactly n attempts — and are counting the number of times you succeed.

Note: In the literature on probability theory, each of the attempts that you make is called a ***Bernoulli trial***.

What This is Called: The ***Binomial distribution with parameters n and p***.

Negative Binomial Distribution

Situations Modelled: One again, you are repeatedly trying to accomplish something — succeeding each time with probability p, where $0 < p < 1$. You stop as soon as you have succeeded k times — and you are counting the number of attempts you must make.

What This is Called: The ***negative Binomial distribution with parameter p (and k)***.

Hypergeometric Distributions

Situations Modelled: Suppose you are sampling from a large population that consists of two groups (so that everyone, in the population that you are sampling from, belongs to exactly one of these groups). In particular, suppose that the population has size N and that K people belong to the first group (so that $N - K$ people belong to the second group).

You sample from the population n times, for a positive integer n and you are interested in the number of times that the person you "sampled" belongs to the first of these groups.

What This is Called: When you are sampling without replacement — so that no person can be sampled more than once — this is called the ***Hypergeometric distribution (with parameters N, K, and n***.

Note: While one could imagine a variant of this where you sample *with* replacement (as suggested in the preparatory material) this would just be a disguised version of the ***Binomial distribution*** — with p set to be K/N.

Approximations

The expressions, used to state probabilities. can become quite complicated. For example, this can happen when the Hypergeometric distribution is used.

When parameters (like the number of trials used) are large, the probabilities that one gets can be numerically close to those that would be obtained using different distributions instead. For example, when sampling from an extremely large population, the probability obtained by sampling without replacement (using the —rather complicated — *Hypergeometric distribution*) can be approximated by sampling with replacement (and working with a — simpler — *Binomial distribution* instead).

When working with *extremely* large populations, you might be interested in the ***limit*** formed, using a probability, as the population size approaches $+\infty$. This leads to ***continuous probability theory***, as mentioned next.

Continuous Probability Theory

This kind of "probability theory" is of interest when the sample space is uncountably infinite. Indeed, the set of all real numbers, and the set of real numbers between some upper bound and lower bound, are all commonly used here.

Definition 1. Let $\Omega = \mathbb{R}$ — or let $\Omega = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ for a pair of real numbers a and b such that $a < b$. An *integrable* function $f : \Omega \to \mathbb{R}$ lis a ***probability density function*** for Ω if it satisfies the following properties.

(a) $f(x) \geq 0$ for every real number x such that $x \in \Omega$.

(b) $\int_{t\in\Omega} f(t)\mathrm{d}t = 1$. That is, if $\Omega = \mathbb{R}$ then

$$\int_{-\infty}^{+\infty} f(t)\mathrm{d}t = 1$$

and if $\Omega = \{x \in \mathbb{R} \mid a \leq t \leq b\}$ then

$$\int_{b}^{a} f(t)\mathrm{d}t = 1.$$

Definition 2. Let Ω be a sample space. A random variable is ***continuous, with density f***, if

$$P(\alpha \leq X \leq \beta) = \int_{\beta}^{\alpha} t \cdot f(t)\,\mathrm{d}t$$

for all real numbers $\alpha, \beta \in \Omega$ such that $\alpha \leq \beta$.

Definition 3. If a random variable $X : \Omega \to \mathbb{R}$ is continuous, wth density $f : \Omega \to \mathbb{R}$, then the ***expected value*** of X is

$$\mathsf{E}[X] = \int_{t \in \Omega} t \cdot f(t)\, dt$$

and the ***variance*** of X is

$$\mathsf{var}(X) = \int_{t \in \Omega} \mathsf{E}[(X - \mathsf{E}[X])^2] = \int_{t \in \Omega} (t - \mu)^2 \cdot f(t)\, \mathrm{d}t$$

where $\mu = \mathsf{E}[X]$.

The expected values and variances of continuous random variables are used in many of the ways that the expected values and variances of discrete random variables are.

One can also consider "classical" *continuous* probability distributions:

- ***Exponential distributions*** are distributions resembling the ones that we get by starting with *geometric distributions* and taking limits (as the number of trials approaches $+\infty$).
- ***Gaussian distributions*** — also called "normal distributions" — are distributions resembling the ones that we get by starting with *Binomial distributions* and taking limits (as the number of trials approaches $+\infty$).