Lecture 13. Continuous Distributions

- In a continuous probability space, the probabilities of events cannot be derived from the probabilities of the individual outcomes in them. In many cases, events containing only finitely many outcomes have zero probability.

- *Example.* What does it mean to pick a point at random from the unit interval, if no point is to be any more probable than any other? If the interval is divided half, then we must grant that the random point must have equal probability of being in either half, and hence the probability of being in the first half must be $1/2$, and the probability of being in the second must be the same. Similarly, if the interval is divided into thirds, the probability of being in the first third must be the same as the probability of being in the second and the same as being in the third, and hence the probability of being in any given third must be $1/3$. In the same manner, if the interval is divided into n subsets each of the same length, then the probability o being in any selected one of the equal parts must be $1/n$. We conclude that the probability of being in a subset of $[0, 1]$, under the assumption that the points are all equivalent in terms of probability, is the total length of the given subset. In particular, the probability of any specific number—say $47/168$—is zero.

- In a continuous probability, the idea of *probability density function* takes the place of the idea of a probability mass function. If the sample space Ω is $\mathbb{R}$ then f is a function on $\mathbb{R}$. If $A \subseteq \mathbb{R}$ is an event, then
$$P(A) = \int_A f(x)dx$$

- We showed previously that if Ω is a discrete sample space (i.e., probability space) with pmf g, and if X is a random variable on Ω, then we can create a new sample space by lumping together the outcomes at which X takes the same value. This new sample space consists of a subset of $\mathbb{R}$ (namely, the set of all possible values of X), with the new pmf f defined by
$$f(x) := P(X = x) = \sum_{X(\omega)=x} g(x).$$
Because we can always make this simplifying move, when we deal with discrete random variables, it is common to assume that the sample space is a subset of $\mathbb{R}$, and that X is simply the identity function on that set.

- *Example.* The binomial random X with parameters n and p has the following pmf:
$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}\ , \ x = 0, 1, 2, \ldots, n.$$
We view $\{0, 1, 2, \ldots, n\}$ as the sample space, and X is simply the identity function. The set of points where X is between 2 and 5 is $\{2, 3, 4, 5\}$. Its probability is
$$P(2 \le X \le 5) = \sum_{x=2}^{5} \binom{n}{x} p^x (1-p)^{n-x}.$$

- Analogously, when we say that X is a continuous random variable, we typically view $\mathbb{R}$ (or a subinterval of $\mathbb{R}$) as the sample space, we think of X as the identity function, and we assume that there is a pdf f on $\mathbb{R}$ that we may integrate to find probabilities:

$$P(a \leq X \leq b) = \int_a^b f(x)\, dx.$$

- *Example.* We say that X is *uniformly distributed* on an interval $[a, b]$ if the probability of being in any subinterval $[c, d]$, with $a \leq c \leq d \leq b$ is equal to $\frac{c-d}{b-a}$. This forces the pdf to be the function

$$f(x) := \begin{cases} \frac{1}{b-a}, & \text{if } a < x < b; \\ 0, & \text{if } x < a \text{ or } b < x. \end{cases}$$

  We have not stipulated values for $f(a)$ and $f(b)$, but they do not matter. The same probabilities are assigned to every interval, regardless of the value of the pdf at these points—or at any finite set of points.