Mathematics 134 – Intensive Calculus for Science 2
Discussion 4 – Distributions and Densities
Thursday March 15, 2001

Background

Last time, we introduced:

1) the *density function* $p(x)$ for a quantity x in a population. This is the function that satisfies the equation: (fraction of the population with $a \leq x \leq b$) $= \int_a^b p(x)\ dx$.
2) The *cumulative distribution function* of x:

$$P(t) = \int_{-\infty}^{t} p(x)\ dx$$

which gives the fraction of the population with $x \leq t$.
3) The *median value of* x: The value $t = T$ such that half the population has $x \leq T$ and the other half has $x \geq T$: $P(T) = \int_{-\infty}^{T} p(x)\ dx = .5$
4) The *mean value of* x: The value

$$\overline{x} = \int_{-\infty}^{\infty} xp(x)\ dx$$

(analogous to the center of mass, since $\int_{-\infty}^{\infty} p(x)\ dx = 1$ here!)

Discussion Questions

A) Give qualitative sketches of graphs of a density function and a cumulative distribution function that could represent the distribution of the quantity $x =$*annual income* in a population where

   1) There is a large middle class
   2) There is a large group of poor people, and small middle and upper classes

B) Recall, last time we introduced the density function for phone call lengths: $p(x) = .4e^{-.4x}$ ($x =$ length in minutes).

   1) What are the median and mean phone call lengths here?
   2) Suppose you selected a random phone call out of the whole "population". What is the *probability* that the length of that call is between 5 and 10 minutes? Explain.

C) Suppose that x is the time (in hours) it takes students in a large calculus lecture class at a state university to finish an exam. Assume that all students finish after 2 hours and that the density function for x is given by

$$p(x) = \begin{cases} x^3/4 & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

1) What fraction of the students take between 1.5 and 2 hours to finish the exam?
2) What is the *median time* to finish the exam?
3) What is the *mean time* to finish the exam?

D) The mean and median values of a quantity x do not have to be equal (although they certainly could be in some cases). Explain why the mean and median can be different in intuitive terms. (Think about means and medians for test scores in a class, for example. Here the density function $p(x)$ would give the fraction of the class earning a score of x. You can picture this with a histogram like our height histogram.)

E) A basic principle of statistics is that (for large enough populations) many quantities have density functions of the form

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}.$$

If so, we say the quantity x is *normally distributed.* The constants σ, μ have the following names here – μ is the *mean value* of x (it is also the median in this case!). σ is another quantity called the *standard deviation.*

1) Plot $p(x)$ above with $\mu = 1, \sigma = 1$, then with $\mu = 1, \sigma = 2$, and finally with $\mu = 1, \sigma = 3$. What happens to the graph as σ increases? Describe in words what the standard deviation of a normal distribution tells you.
2) Let $\mu = 1, \sigma = 1$ as in your first plot above. Sketch the cumulative distribution function for this density function.
3) The distribution of IQ scores for people is modeled by a normal distribution with $\mu = 100$, and $\sigma = 15$. Write the integral giving the fraction of the population with IQ scores between 115 and 120. Approximate this fraction using a LHS with $n = 10$.

Assignment

Group write-ups due in class on Tuesday, March 20.