Avani Patel
Nicole Nagel

Monday, April 7th Class Notes

At the beginning, we went over administration stuff involving projects. We have a homework due on Monday, April 14th.

Next, we did some review of last class. We reviewed probability mass tables for discrete random variables. Discrete random variables have a set number of outcomes, unlike continuous variables. You can make continuous variables discrete by assigning them to buckets or bins, giving them ranges.

One of the equations we went over is the weighted average of possible outcomes. This is also the expected value, or the average outcome of a random variable:

$$E(X) = \sum_{k=1}^{D} W * x$$

We then calculated the expected value of a function of random variables:

$$E(Y) = E\left(f(x)\right) = \sum_{k=1}^{D} f(x) * W$$

Next, we learned about Jensen's Inequality, which is as follows:

In general (except for linear functions), $E\big(f(x)\big) \neq f(E(x))$.
We went over a short proof for why this works for linear equations, (although it is not vital to know). Given $f(x) = ax + b$:

$E\big(f(x)\big) = \mathrm{E(ax\ +\ b)} = \mathrm{a * E(x)} + b = f\big(E(x)\big).$

Here is an example proving Jensen's inequality, where $f(x)$, or the quantity demanded, is $Y = 1000 * x^{-2}$

X_k (price)	W_k (probability)
2	0.2
3	0.5
4	0.3

$$E\big(f(x)\big) = 0.2 * 1000 * 2^{-2} + 0.5 * 1000 * 2^{-2} + 0.3 * 1000 * 2^{-2} = 124.306$$
$$f\big(E(x)\big) = f(2 * 0.2 + 3 * 0.5 + 4 * 0.3) = f(3.1) = 104.058$$

As you can see above, $E\big(f(x)\big) \neq E(x)$ for nonlinear functions.

Next, we discussed variance, which is the measure of how spread out, or how uncertain you are, about a random variable. Variance is calculated as follows:

$$var(x) = E[\big(x - E(x)\big)^2]$$

Where $E(X)$ is the expected value. This can be simplified:

$$= E[(x - \mu)^2]$$
$$= E\big(f(x)\big), where\ f(x) = (x - E(x))^2$$

Now, we discussed the difference between probability mass functions (PMFs) and probability density functions (PDFs).

PMFs are discrete, and usually are used when a probability mass table would be too long to list. The equation for PMF is:

$$P(X = x_k) = f(x_k)$$

For example,

X = the number of no-shows on a flight from Austin to Dallas.
You can model X using a binomial distribution (this is the most simple form of PMF):

$$P(X = x_k) = \binom{N}{x_k} p^{xk}(1 - p)^{N - xk}$$

N = # seats on flight
P = probability that an individual person is a no-show.

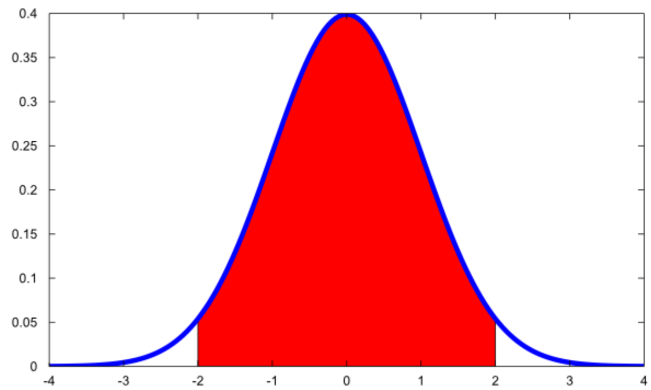
$$\binom{N}{x_k} = binomial\ coefficient$$

The binomial coefficient is the number of ways in which x_k people could fail to show up.

There are many other PMFs, but we will focus on binomial distributions. Another example is the Poisson distribution, which is used for measuring counts (ex. how many calls to a call center).

Probability Density Functions, PDFs, are for continuous, random variables. (Ex. price.)

Here is a PDF of a normal distribution:



On this graph, the red area is the probability that x will fall between -2 and 2.

$$P(-2 \leq x \leq 2) = \ area\ under\ PDF$$

Now, we are moving on to joint distributions.

Consider this example:

X and Y are both random variables.

x = price of AAPL
y = value of a 30year treasury bond with a specific maturity

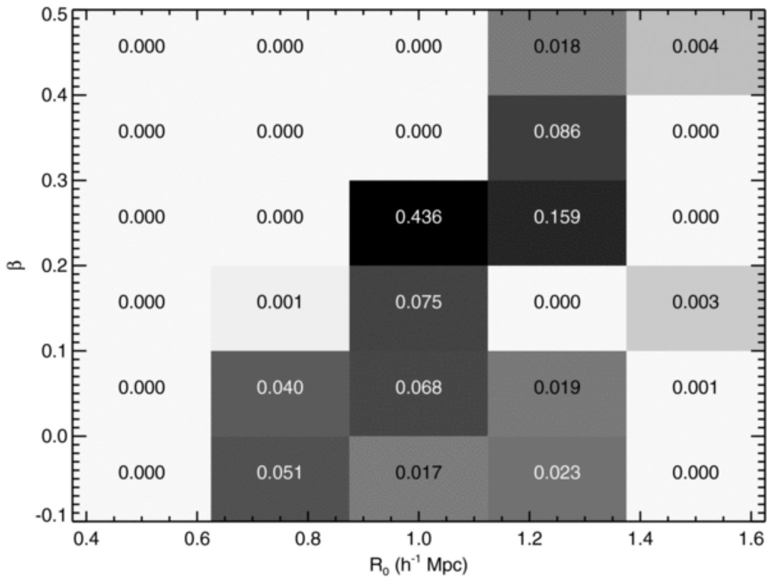
$$f(x, y) = \ a\ policy\ or\ a\ decision$$

We evaluate $f(x, y)$ to see if our decision or policy is good, bad, or in-between.

For joint distributions, you must list every possible combination of x and y. These tables can then get very long if you have even just 10 possibilities for x and 10 for y (this would be 100 rows…ugh).

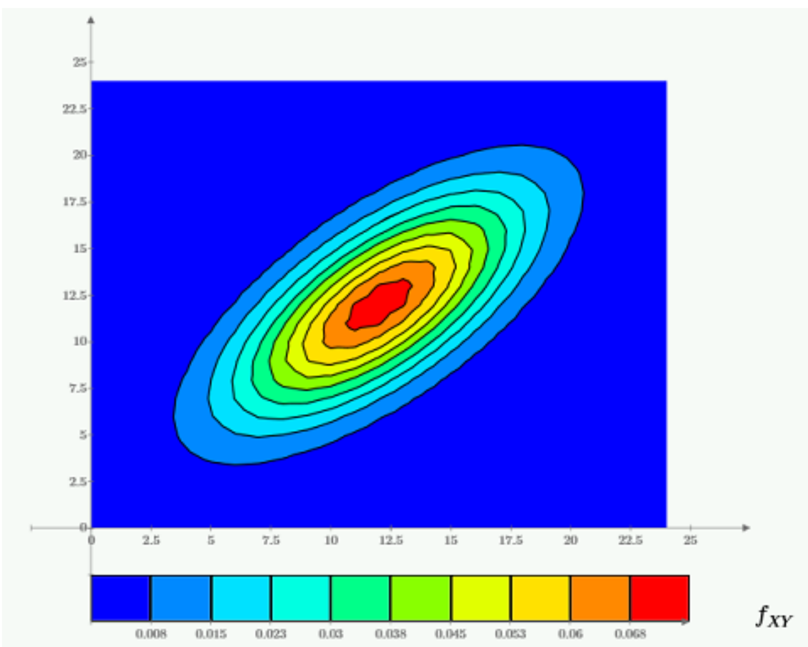
A table isn't the best way to visually represent this data, so we went over several ways to do so, like gray scaled point clouds and boxes, and contour plots.

Here is an example of a joint distribution table coded in gray-scale based on probability density:



As this is made more precise, these blocks become smaller gray-scale points, and turn into a point cloud.

Here is an example of a contour plot. Each ring has the same probability:



Our final concept for the day is covariance. The equation for covariance is:

$$cov(x, y) = E\{[x - E(x)] * [y - E(y)]\}$$
$$E\{f(x, y)\} where\ f(x, y) = [x - E(x)] * [y - E(y)]$$

A strategy for covariance is to apply the function to every outcome, and then sum the weighted function values.

$$cov(x, y) = \sum_{k=1}^{D} [x_k - E(x)] * [y_k - E(y)] * w_k$$

We did a quantitative example of covariance in Excel with the CSV linked to the week 12 page. (Note that the csv file was updated after Monday's class with new numbers that work better.)

The top few rows and the bottom row (which includes the necessary sums) of the Excel file is pasted below:

x	y	JointProb	w_k * x_k	w_k * y_k	x-difference	y-difference	covariance
-2.5	-2.5	0.00158754	-0.00396885	-0.00396885	-3.45318407	-2.97669621	0.016318451
-2.5	-1.5	0.00711487	-0.017787175	-0.010672305	-3.45318407	-1.97669621	0.048565362
-2.5	-0.5	0.00431539	-0.010788475	-0.002157695	-3.45318407	-0.97669621	0.014554567
			sum: 0.95318407	sum: 0.47669621			sum: 0.891451086 = total covariance

To get column 4 and 5 values, multiply x by JointProb, then y by JointProb, respectively. The sums of these columns is used to calculate the x-difference and y-difference.
x-difference is x – sum(w_k * x_k)
y-difference is y – sum(w_k * y_k)

Covariance is calculated by multiplying JointProb * x-difference * y-difference, then summed for the total covariance.