MML - 6.2 Disc & Cont Probs

When target space T is discrete we specify the probability a RV X takes a particular value $x \in T$
- $P(X = x)$

This is known as the Probability Mass Func

When Target space T is continuous. It is more natural to specify an interval for the prob of a random variable
- $P(a \leq X \leq b)$ for $a < b$

known as cumlative distribution

Note Univariate = dist of single RV

dists of more than one variable = Multivariate
& consider a vector of random variables

6.21 Discrete Probabilities

With Discrete targets, the joint probability is the cartesian prod of the target space of each the random variables

Joint prob is the entry of both values joints

$$P(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$$

Attrib label

n_{ij} = num of events with state x_i & y_i
N = total num of events

Joint Prob = intersect of both Probs

$$P(X = x_i, Y = y_j) = P(X = x_i \cap Y = y_j)$$

Lazily written as $P(x, y)$

The marginal Prob we see x_i regardless of y_i = $P(x)$
Conditional Prob of x = $P(y|x)$

Example 6.2 - Discrete

RV's = X^5, Y^3

n_{ij} = num events w/ $X = x_i$ $Y = y_j$
N = total events

C_i = Sum of individ freqs for the i^{th} column
- $C_i = \sum_{j=1}^{3} n_{ij}$ (how many of outcomes per X_i)

R_j = Sum of individ freqs for the i^{th} column
- $R_j = \sum_{i=1}^{5} n_{ij}$ (how many X_i values per Y_i)

$$P(X = x_i) = \frac{C_i}{N} = \frac{\sum_{j=1}^{3} n_{ij}}{N}$$

$$P(Y = y_i) = \frac{C_j}{N} = \frac{\sum_{i=1}^{5} n_{ij}}{N}$$

by convention Probs must sum to one:

$$\sum_{i=1}^{3} P(X = x_i) = 1 \quad \text{or} \quad \sum_{j=1}^{3} P(Y = y_j) = 1$$

the conditional prob is given as

$$P(Y=y_j \mid X=x_i) = \frac{n_{ij}}{c_i}$$

$$P(X=x_i \mid Y=y_j) = \frac{n_{ij}}{r_j}$$

6.2.2 Continuous Probs

Real valued Random Variables