Random Variables

A random variable is a function defined on the outcomes in a sample space.

Ex. Toss a coin 10 times. The random variable X is the number of heads in 10 tosses.

Ex. In a class of students, X is a randomly selected student's exam score.

All our random variables will be defined on finite sample spaces.

For each value x that X takes on, the set of outcomes for which X=x is an event.

We write P(X=x) for the probability of this event. When we list the values of X and their corresponding probabilities, that is the probability distribution of X.

Ex. Toss a fair coin 5 times. X is the number of heads in 5 tosses.

The event "X=0" is {TTTTT}

The event "X=1" is {HTTTT, THTTT, TTHTT, TTTHT,TTTTH}

S is all the 5-long sequences of H and T. n(S)=32 so each outcome has probability 1/32

The event "X=k" has C(5,k) different outcomes and so has probability C(5,k)/32

The probability distribution of X is

k	0	1	2	3	4	5
P(X=k)	1/32	5/32	10/32	10/32	5/32	1/32

The probabilities of other events can be found by adding probabilities:

P(X<3)=(1+5+10)/32 P(2<X<5)=(10+5)/32

Ex. A student takes 12 credits and receives the following grades.

Course grade	A	B	C	C	D	F
# credits	3.0	3.0	2.0	1.0	2.0	1.0

For S={A, B, C, D, F} the random variable X is the number of points assigned to the letter grade in a 4-pt grading system. X(A)=4 X(B)=3 X(C)=2 X(D)=1 X(F)=0

In this semester for this student the distribution of X is

k	4	3	2	1	0
P(X=k)	3/12	3/12	3/12	2/12	1/12

Now we can define the expected value of X and find the student's gpa.

Expected Value: If X is a random variable with values $x_1, x_2, ..., x_n$ and respective probabilities
$p_1, p_2, ..., p_n$ then the expected value of X is E(X)= $p_1 x_1 + p_2 x_2 + ... + p_n x_n$

Ex. The student's gpa is the expected value of X, E(X)=4x3/12 + 3x3/12 + 2x3/12 + 1x2/12 + 0x1/12 = 25/12=2..08333…

Expected value is a weighted average of the values of X with the weights provided by the probabilities.

Ex. If X is the number of heads in 5 tosses of a fair coin then the expected value of X is E(X)=0x1/32 + 1x5/32 + 2x10/32 + 3x10/32 + 4x5/32 + 5x1/32 = 2.5

Ex. A person invests $30,000, $40,000, and $50,000 in projects A, B, and C respectively. The returns for projects A, B, and C are 6%, 12% and 15% respectively. What is the average return on the money?
A total of $120,000 is invested. If we think of the investment amount as the frequency and the amount divided by 120,000 as the relative frequency we have

Return	.06	.12	.15
Rel. freq	1/4	1/3	5/12

The weighted average is .06x1/4 + .12x1/3 + .15x5/12 = .1175 or 11.75%

Ex. This is like some problems in the book:
An insurance company issues a 1 year term life insurance policy which pays the beneficiary $20,000 if the person dies within a year. The company's tables show the person has a .15 chance of dying within the year. At what price for the policy does the company break even on average?
Let X be the company's payout. X=0 if the holder lives and X=$20,000 if the holder dies.
P(X=0)=1-.15=.85 and P(X=20,000)=.15   E(X) represents the company's average payout and E(X)=0x.85 + 20,000x.15 = 3000. They must charge at least $3000 for such policies to break even.

Ex.: A student has a 3.6 gpa and 85 credits. He wants to guess his grades and figure his new gpa. He estimates his grade distribution for the new semester to be

Grade	A	B	C	D	F
#credits	2	4	3	3	0

We want to figure the sum of grade pointsx#credits divided by the total number of credits.

This is (3.6x85 + 4x2 + 3x4 + 2x3 + 1x3)/ (85 + 12) =3.4536…

A new subject, Odds:

If E is an event in a sample space then the odds in favor of E are $\frac{P(E)}{P(E^C)}$.

If the odds in favor of E are $\frac{a}{b}$ then P(E)= $\frac{a}{a+b}$

Ex. The weather report predicts a 60% chance of rain. What are the odds it will rain?
P(rain)=.6   P(no rain)=.4 so the odds are 6/4 or 3/2.

Ex. The odds in favor of a certain team winning a baseball game are 2/3. Then the probability this team will win is 2/(2+3) = 2/5.

Standard Deviation and Variance:

If X is a random variable, the variance of X is Var(X)=E($(X-\mu)^2$) where μ =E(X). Var(X) is the average squared distance to the mean.

The standard deviation of X is $\sigma(X)=\sqrt{Var(X)}$.

A sometimes more convenient expression for Var(X) is E(X^2) - μ^2 .

The standard deviation is a measure of the spread of the distribution. If two distributions have the same expected value, or mean, the one that has higher probability of being farther from the mean has the higher standard deviation.

Ex. Compute the expected value, variance and standard deviation for the distribution shown.

x	0	1	2
P(X=x)	1/2	1/3	1/6

E(X)= μ =0x1/2 + 1x1/3 + 2x1/6 = 2/3

E(X^2)= 0x1/2 + 1x1/3 + 4x1/6 = 1

Var(X) = 1-4/9 = 5/9

$\sigma(X)=\sqrt{\frac{5}{9}}$ =.74533…

Ex. x	0	1	2
P(X=x)	1/3	2/3	0

E(X)=2/3

E(X^2) =2/3

Var(X)=2/3-4/9 = 2/9

$\sigma(X)=\sqrt{\frac{2}{9}}$ =.4714…This distribution is less spread out than the one above which has the same mean.

Chebychev's Inequality: The probability that X is within k standard deviations of the mean is at least 1-$\frac{1}{k^2}$. P($\mu-k\sigma$ $\leq$X$\leq$ $\mu+k\sigma$) $\geq$1-$\frac{1}{k^2}$

Ex. #28 from Tan:

A Christmas tree light has an expected life of 200 hr. and a standard deviation of 2 hr.

a) Estimate the probability that one of these lights will last between 190 and 210 hrs.

Find k: How many standard deviations from the mean=200 are 190 and 210? The standard deviation is 2 and 190=200-5x2, 210=200+5x2 so k=5.

So the probability that X lies between 190 and 210 is at least 1-$\frac{1}{5^2}$ =.96

b)Suppose 150,000 of these lights are used by a large city as part of its Christmas decorations. Estimate the number of lights that will need replacement between 180 and 220 hr of use.
Estimate the probability that a light will last between 180 and 220 hrs. Find k:
180 = 200-20=200-10x2 220=200+10x2 so k=10. The probability is at least 1-1/100=.99

At least .99x150000=148,500 lights expected to burn out between 180 and 220 hrs.

Ex. #30 from Tan:
Sugar packaged by a certain machine has a mean weight of 5 lb and a standard deviation of .02 lb. For what value of c can the manufacturer of the machinery claim that the sugar packaged by this machine has a weight between 5-c and 5+c lb with probability at least .96?

Find k: This time we are told the probability is at least .96 so $1-\frac{1}{k^2}=.96$.

$\frac{1}{k^2}=.04$ and $k^2=1/.04=25$ so k=5. Then $c=5\text{x}\sigma=5\text{x}.02=.1$