```
Math 55 - Spring 2004 - Lecture notes #20 - Apr 8 (Tuesday)

Goals for today:    random variables
                    expection (average, mean) of random variables

 DEF: Let S be the sample space of a given experiment, with probability
      function P. A _random variable_ is a function f:S -> Reals.

 EX: Flip a biased coin once, S1 = {H,T}, P1(H) = p,
     f1(x) = {1 if x=H, -1 if x=T}
     f1 = amount you win (if f1>0) (or lose, if f1<0) if you bet $1 on H.

 EX: Flip a biased coin n times, S2 = {all sequences of H, T of length n}
ASK&WAIT: What is P2(x), if x has i Heads?
       f2(x) = #H - #T = #H - (n-#H) = 2*#H - n
       f2 = amount you win (or lose) if you bet $1 on H on each flip

 EX: Let S3 = result of rolling a die once, P3(any face) = 1/6
     Let f3(outcome e) = value on top of die (an integer from 1 to 6)
 EX: Let S4 = result of rolling a pair of red and blue dice 24 times
          = { ((1,1),(1,1),...,(1,1)),...,((6,6),(6,6),...,(6,6))}
              <----- 24 times ------>     <----- 24 times ------>
ASK&WAIT:   What is P4(any outcome x in S4)?
     Let f4(outcome x in S4) = { +1 if a pair of sixes appears in e }
                               { -1 otherwise                      }
     We can interpret f4 as the amount of money we win (or lose) by
     betting on getting a pair of sixes

 EX: S5 = {US population}, P5(person x in S) = 1/|S5|,
     Let f5(person x in S) = { +1 if x has a particular disease }
                             {  0 if x does not                 }

 EX: S6 = {all permutations of 1 to n}, P6(any permutation) = 1/n!
     f6(any permutation x) = time for your sorting algorithm to sort x

EX: Suppose you flip a fair coin, and win $1 if it comes up H,
    lose $1 if it comes up T
ASK&WAIT: What is the "average" amount you expect to win after N flips?

 DEF: Given S, P and random variable f, the _Expected Value_
      (also called Mean or Average) of f is
          E(f) = sum_{all outcomes x in S} P(x)*f(x)
```

This is the "average" value of f ones gets if one repeats the experiment
a great number of times.
EX: With S1, P1, f1 as before,
E(f1) = (+1)*(p) + (-1)*(1-p) = 2*p-1
= 0 if coin fair (p=1/2)
Imagine betting $1 on getting H. Then E(f1) is the amount you expect to
win (if E(f1)>0) or lose (E(f1)<0) on the bet. If E(f1)=0, you break even

EX: With S2, P2 and f2 as before,
If we flip a coin N times, we expect E(f2) to be the amount we win
betting $1 on flip to get H; and intuititively this should be
N*E(f1) = N*(2*p-1)
Formally, we get
E(f2) = sum_{sequences x of n Hs and Ts} f2(x)*P2(x)
= sum_{sequences x of n Hs and Ts} (#H in x)*P2(x)
looks complicated, but later we will see that our intuition was right,
and there is an easier way to do it that matches our intuitive approach

EX: With S3, P3, f3 as before,
E(f3) = (1/6)*1 + (1/6)*2 + ... + (1/6)*6 = 21/6 = 7/2

EX: With S5, P5, f5 as before,
E(f5) = sum_{persons x} f5(x)*P5(x)
= sum_{sick persons x} f5(x)*P5(x) + sum_{healthy persons x} f5(x)*P5(x)
= sum_{sick persons x} 1*(1/|S5|) + sum_{healthy persons x} 0*(1/|S5|)
= P(random person is sick)

EX: S6, P6, f6 as before, then E(f6) = average time for your algorithm to sort

EX: With S4, P4, f4, seem like you need to sum over all 6^48 sequences,
We need a simpler way:

DEF: P(f=r) = sum_{all outcomes x in S such that f(x)=r} P(x)

EX: With S1, P1 and f1 as before
P1(f1=1) = P1(H) = p, P1(f1=-1) = P1(T) = 1-p
EX: With S2, P2 and f2 as before
ASK&WAIT: What is P2(f2=i)?

EX: With S3, P3 and f3 as before
P3(f3=k) = 1/6 for k=1,2,...,6 and P3(f3=k)=0 otherwise

```
 EX:  With S4, P4 and f4 as before,
      P4(f4=1) = sum_{all outcomes x in which a pair of sixes appears} P4(x)
               = P4(a pair of sixes appears)
ASK&&WAIT: What is P4(f4=-1)?

 EX:  With S5, P5 and f5 as above,
ASK&WAIT: what is P5(f5=1)? P5(f5=0)?

 Thm: E(f) = sum_{numbers r in range of f} r*P(f=r)
    Proof: Write down proof for S finite, but same for S countable
           Let {r1,r2,...,rk} be numbers in range of f, and write
           S = S1 U S2 U ... U Sk where
           Si = {outcomes x in S such that f(x)=ri}
           and so P(Si) = P(f=ri)
           Note that all Si are pairwise disjoint, so we can write
           E(f) = sum_{x in S} f(x)*P(x)
                = sum_{x in S1} f(x)*P(x) + sum_{x in S2} f(x)*P(x)
                  + ... + sum_{x in Sk} f(x)*P(x)
                = sum_{x in S1} r1*P(x) + sum_{x in S2} r2*P(x)
                  + ... + sum_{x in Sk} rk*P(x)
           Look at one term:
           sum_{x in Si} ri*P(x) = ri * sum_{x in Si} P(x)
                                 = ri * P(Si)
                                 = ri * P(f=ri)
           so E(f) = r1*P(f=r1) + r2*P(f=r2) + ... + rk*P(f=r3)
                   = sum_{number r in range of f} r*P(f=r)
           as desired.

 EX: With S3, P3 and f3 as above,
     E(f3) = sum_{k=1 to 6} k*P(f=k) = sum_{k=1 to 6} k*(1/6) = 7/2 as before

 EX: With S4, P4, f4 as above,
     E(f4) is the average amount one wins (if E(f4)>0) or loses (if E(f4)<0)
     every time one plays.
     E(f4) = sum_{numbers r in range of f} r*P(f4=r)
           = +1*P4(getting pair of sixes) + (-1)*P4(not getting pair of sixes)
           = P4(getting pair of sixes) - P4(not getting pair of sixes)
ASK&WAIT:     What is P4(not getting pair of sixes)?
     P4(getting pair of sixes) = 1 - P4(not getting pait of sixes)
                               ~ 1-.5086 = .4914
     and E(f4) = .4914 - .5086 = -.0172, so you lose in the long run
```

```
    Note: In 1654 the gambler Gombaud asked Fermat and Pascal whether
          this was a good bet, inadvertently starting the field of
          probability theory
    Note: If we do 25 rolls instead of 24,
          P4(not getting a pair of sixes) drops to (35/36)^25 ~ .4945
          P4(getting pair of sixes) grows to .5055, so it is a good bet.

 EX: Let S5, P5, f5 be as above. Then
    E(f5) = (+1)*P(f5=1) + 0*P(f5=0)
          = P(f5=1) = P(person sick)
    This is a special case of the following lemma:

 Lemma: Let S be a sample space, E subset S any event, and
       f(x) = {1 if x in E     }
              {0 if x not in E }
       Then E(f) = P(E)
ASK&WAIT: proof?

EX: S2, P2, f2 as above:
    E(f2) = expected win betting $1 on a coin N times
          = sum_{i=0 to N} i*P2(getting i heads)
          = sum_{i=0 to N} i*C(N,i)*p^i^(1-p)^(N-i)
    still isn't simple, so need a new idea:

 Thm: Let S and P be a sample space and probability function, and
      let f and g be two random variables. Then
           E(f+g) = E(f) + E(g)
      Proof: Let h=f+g be a new random variable.
           Then E(h) = sum_{outcomes x in S} h(x)*P(x)
                     = sum_{outcomes x in S} (f(x)+g(x))*P(x)
                     = sum_{x} f(x)*P(x) + sum_{x} g(x)*P(x)
                     = E(f)              + E(g)

  Corollary: Let S and P be as above, and h = f1 + f2 + ... + fn
      Then E(h) = E(f1) + E(f2) + ... + E(fn)

  EX: Let S2, P2, f2 be as before. Then we can write
      f2 = g1 + g2 + ... + gN where
      gi(x) = { +1 if i-th flip = H }
              { -1 if i-th flip = T }
      and E(f2) = E(g1) + E(g2) + ... + E(gN)
```

```
     For any i E(gi) = (+1)*P(H) + (-1)*P(T) = p - (1-p) = 2*p-1
     so E(f2) = N*(2*p-1)
     which matches our original intuition about making N independent
     bets in a row (whew!)

 EX: Let S4, P4, and f4 be as before. Suppose you also make the side bet
     that you win 2 if at least 8 fives come up, and lose 2.5 if
     fewer than 8 fives come up. Is this joint bet worth making?
     Answer: Let g(x) = { +2 if at least 8 fives come up in x  }
                        { -2.5 if at most 7 fives come up in x }
     P(g=+2) = P(at least 8 fives)
             = sum_{i=8 to 48} C(48,i) * (1/6)^i * (5/6)^(48-i)
             ~ .55992
     P(g=-2.5) = P(at most 7 fives)
               = 1 - P(at least 8 fives)
               = 1 - .55992 = .44008
     E(g) ~ +2*.55992 - 2.5*.44008 ~ .0196
     Then the value of the joint bet f4+g is
            E(f4+g) = E(f4)+E(g) ~ -.0172+.0196 = .0024
     and being positive, is worth making.

EX: Suppose you shoot at a target, and miss it with probability p each
    time you try. What is the expected number of times you have to try
    before getting a hit?
    S = { H, MH, MMH, MMMH, .... }
    P( MM...MH ) = p^#M * (1-p)
    f( MM...MH ) = #shots = #M + 1
    We want E(f) = sum_{m=0}^infinity (m+1)*p^m*(1-p)
         Recal     sum_{m=0}^infinity p^m = 1/(1-p)
         so d/dp ( sum_{m=0}^infinity p^m ) = d/dp ( 1/(1-p) )
         or    sum_{m=0}^infinity m*p*{m-1} = 1/(1-p)^2
         or    sum_{m=0}^infinity m*p^m*(1-p) = p/(1-p)
            so sum_{m=0}^infinity (m+1)*p^m*(1-p) =
                         p/(1-p) + (1-p)/(1-p) = 1/(1-p)
         so E(f) = 1/P(hit)
    So if P(M)=.99, you need to take 1/(1-.99) = 100 shots on average to hit

EX: Suppose homework from 350 students is collected, graded, randomly
    shuffled, and handed back. What is the expected number of students
    who get their own homework back?
    S = {permutations of 1 to 350}, P(any permutation) = 1/350! ~ 8e-741
    f(permutation x) = #homeworks returned to right students,
```

```
    We want E(f)
    Let fi(x) = { 1 if student i gets right homework back }
                { 0 otherwise }
    Then f(x) = f1(x) + f2(x) + ... + f350(x)
    and E(f) = E(f1) + ... + E(f350)
    Now E(fi) = P(student i gets right homework)
              = (# permutations where student i gets right homework)/350!
              = (# permutations of other 349 homeworks)/350!
              = 349! / 350! = 1/350
    so E(f) = 350*(1/350) = 1
    Result would be true for any number of students!

EX: Recall definition of independent sets:
    P(A inter B) = P(A) * P(B); intuitively, this means that knowing whether
    or not you are a member of A tells you nothing about whether you are
    a member of B

EX: S = { 2 coin flips } = { HH, HT, TH, TT } with P(H)=p, P(T)=q=1-p
    A = { HH, HT }, B = { HH, TH }
    Then P(A inter B) = P(HH) = p^2 = P(A)*P(B)
    Let f1(e) = { 1 if first coin H     f2(e) = { 1 if second coin H
                { 0 otherwise                   { 0 otherwise
    Then P(A inter B) = P(f1=1 and f2=1) = P(A)*P(B) = P(f1=1)*P(f2=1)
    Can also check that both A and complement(A) are independent with
    B and complement(B), i.e.
        P(f1=r1 and f2=r2) = P(f1=r1)*P(f2=r2) for r1, r2 = 0, 1
    (this was homework due yesterday!)

DEF Let f, g be random variables. Then we call
       f and g _independent_ if for all values of r and s
       P({x: f(x)=r and g(x)=s}) = P({x: f(x)=r}) * P({x: g(x)=s})

    This generalizes situation where f,g= 1 or 0 only. It still means
    intuitively that knowing about the value of f says nothing about
    the value of g. In this case, the result P(A inter B)=P(A)*P(B)
    generalizes as follows:

 Thm 1: If f and g are independent, then E(f*g)=E(f)*E(g)
  proof: E(f*g) = sum_{x in S} f(x)*g(x)*P(x)
                = sum_{pairs (r,s)} sum_{x such that f(x)=r and g(x)=s}
                              f(x)*g(x)*P(x)
                     because we do the same sum over x, just grouping
```

```
                 together those x where f(x)=r and g(x)=s
             = sum_{pairs (r,s)} sum_{x such that f(x)=r and g(x)=s} r*s*P(x)
             = sum_{pairs (r,s)} r*s* sum_{x such that f(x)=r and g(x)=s} P(x)
             = sum_{pairs (r,s)} r*s* P(f=r and g=s)
             = sum_{pairs (r,s)} r*s* P(f=r)*P(g=s)
                 by independence
             = sum_{r} sum_{s} r*s* P(f=r)*P(g=s)
                 just a different way of summing over all pairs (r,s)
             = sum_{r} r*P(f=r) * sum_{s} s*P(g=s)
             = E(f)             * E(g)
    as desired.

EX: Flip a biased coin 10 times
    S = {all sequences of 10 Hs and Ts}, P(x) = 1/2^10
    Let f = #H - #T in first 5 flips
    Let g = "turn last 5 flips into binary number b, via 1=H and 0=T"
    Then f and g are independent, because they depend on independent
      events (first 5 flips vs last 5 flips) so E(f*g)=E(f)*E(g)
    What are E(f)? E(g)? E(f*g)?
        E(f) = E(#H) - E(#T) = 5*p - 5*(1-p) = 10*p-5
        To compute E(g), let b4 = { 1 if coin 6 = H, 0 otherwise }
                             b3 = { 1 if coin 7 = H, 0 otherwise }
                             ...
                             b0 = { 1 if coin 10 = H, 0 otherwise }
        Then b = b4*2^4 + b3*2^3 + b2*2^2 + b1*2 + b0
        so E(g) = E(b) = E(b4)*2^4 + ... + E(b0)
                       = p*2^4 + ...     + p
                       = p*31
        Finally, E(f*g)=E(f)*E(g)=(10*p-5)*p*31
```