Lecture 18: given a random variable $X$ and its probability distribution $P(X = x)$ we may consider three numbers:

1. The mean, expected value, $\mu = E(X) = \sum_{\text{all } x} x P(X = x)$
   (this is the quantity that is easier to treat mathematically.)

2. Mode: $\text{mode}(X) = k$ for which $P(X = k)$ is maximum
   $\downarrow$
   $= \arg \max_k P(X = k)$  (may be not unique)

3. Median: $m(X)$ is a number $m$ for which
   $P(X < m) \geq \frac{1}{2}$ and $P(X \leq m) \geq \frac{1}{2}$
   (may be not unique).

   - Example: for the Binomial$(n, p)$ distribution,
     with $n$ large and $p$ not close to 0 or 1,
     they coincide. In fact the distribution is symmetric around $np$.

Remember the geometric distribution?

Repeated, independent trials

\( p \) : probability of success
\( q = 1 - p \) : probability of failure
\( X \) : the number of trials until first success

\[ P(X = k) = q^{k-1} p, \quad k = 1, 2, 3, 4, \ldots \quad (k \in \mathbb{N}) \]

Note: \[ \sum_{k=1}^{\infty} p(X = k) = p \sum_{k=1}^{\infty} q^{k-1} = p \left( 1 + q + q^2 + q^3 + \ldots \right) \]
\[ = p \sum_{j=0}^{\infty} q^j = \frac{p}{1-q} = \frac{p}{p} = 1 \]
Mean:

\[ E(X) = \sum_{k=1}^{\infty} k q^{k-1} p = p \sum_{k=1}^{\infty} k q^{k-1} = \]
\[ = p \sum_{k=1}^{\infty} \frac{d}{dq} q^k = p \frac{d}{dq} \left( \sum_{k=1}^{\infty} q^k \right) \]
\[ = p \frac{d}{dq} \left( \frac{1}{1-q} - 1 \right) = p \frac{d}{dq} \left( \frac{q}{1-q} \right) = p \frac{1-q+q}{(1-q)^2} = p \frac{1}{p^2} \]

\[ \Rightarrow E(X) = \frac{1}{p} \]

Mode = 1 (maximum at \( x = 1 \)).

mode = 1 < median < mean

 Median: Hard to compute!

\[ P(X \leq m) = \sum_{k=1}^{m} k q^{k-1} p = p \sum_{k=1}^{m} q^{k-1} = p \frac{1-q^m}{1-q} = 1-q^m \]

\[ S_m \]

\[ S_m = 1 + q + q^2 + \ldots + q^{m-1} \]
\[ q S_m = q + q^2 + q^3 + \ldots + q^m \]
\[ \Rightarrow (q-1) S_m = q^m - 1 \]
\[ S_m = \frac{q^m - 1}{q-1} = \frac{1-q^m}{1-q} \]

Set \( 1-q^m = \frac{1}{2} \) \( \Rightarrow q^m = 0.5 \)

\[ m \log q = \log \frac{1}{2} = -1 \] \( \Rightarrow \)
\[ m = \frac{-1}{\log(1-p)} \]

Since we're looking for smallest \( m \) for which \( P(X \geq m) > \frac{1}{2} \)

\[ m = \left[ \frac{-1}{\log(1-p)} \right] \]

where \( \lceil x \rceil = \text{smallest integer} > x \).

Remark: \( p = \frac{1}{2} \) \( \Rightarrow \) \( E(X) = 2, \) \( m = 1 \).
Theorem: given two random variables $X$ and $Y$, we have
\[ E(X+Y) = E(X) + E(Y). \] (*)

(Interpretation: if I play two games and 
$X =$ money I win from first game, $E(X) =$ expected ... 
$Y =$ money I win from second game, $E(Y) =$ expected ... 
then expected win from both games = sum of expected wins of 
the two games if I played them separately.)

Remark: (*) is always true, even if $X$ and $Y$ are not independent. 
We shall see that $E(XY) \neq E(X)E(Y)$ in general, 
but if $X, Y$ independent then

Proof of (*) : Let $S = X + Y$. $E(S) =$ ? We have 
\[ P(S = k) = \sum_{(x,y): x+y = k} P(x,y), \text{ where} \]
\[ P(x,y) = P(X = x, Y = y). \]

So $E(X+Y) = E(S) = \sum_{k \in \mathbb{K}} k P(S = k)$
\[ = \sum_{k \in \mathbb{K}} k \sum_{(x,y): x+y = k} P(x,y) = \sum_{k \in \mathbb{K}} \sum_{(x,y): x+y = k} k P(x,y) \]
\[ = \sum_{k \in \mathbb{K}} \sum_{(x,y): x+y = k} (x+y) P(x,y) \]
\[ = \sum_{x \in \mathbb{X}} \sum_{y \in \mathbb{Y}} [x P(x,y) + y P(x,y)] \]
\[ = \sum_{x \in \mathbb{X}} \sum_{y \in \mathbb{Y}} x P(x,y) + \sum_{y \in \mathbb{Y}} \sum_{x \in \mathbb{X}} y P(x,y) \]
\[ = \sum_{x \in \mathbb{X}} x P(X = x) + \sum_{y \in \mathbb{Y}} y P(Y = y) = E(X) + E(Y). \]
Example: I roll a die twice.

\[ X_1 = \text{number shown first time} \]
\[ X_2 = \text{number shown second time} \]
\[ S = X_1 + X_2. \quad E(S) = ? \]

"Brute force" method: compute \( P(S = k) \) first, then \( E(S) = \sum_k k P(S = k) \).

Possible ordered pairs (equally likely outcomes):

- \((1,1)\)
- \((1,2)\)
- \((1,3)\)
- \((1,4)\)
- \((1,5)\)
- \((1,6)\)
- \((2,1)\)
- \((2,2)\)
- \((2,3)\)
- \((2,4)\)
- \((2,5)\)
- \((2,6)\)
- \((3,1)\)
- \((3,2)\)
- \((3,3)\)
- \((3,4)\)
- \((3,5)\)
- \((3,6)\)
- \((4,1)\)
- \((4,2)\)
- \((4,3)\)
- \((4,4)\)
- \((4,5)\)
- \((4,6)\)
- \((5,1)\)
- \((5,2)\)
- \((5,3)\)
- \((5,4)\)
- \((5,5)\)
- \((5,6)\)
- \((6,1)\)
- \((6,2)\)
- \((6,3)\)
- \((6,4)\)
- \((6,5)\)
- \((6,6)\)

Range \( S \) = \{2, 3, ..., 12\}

- \( P(S = 2) = P(X_1 = 1, X_2 = 1) = \frac{1}{36} \)
- \( P(S = 3) = P(2,1) + P(1,2) = \frac{2}{36} \)
- \( P(S = 4) = P(3,1) + P(2,2) + P(1,3) = \frac{3}{36} \)

etc. ... Will get, after long procedure: \( E(S) = 7 \)

"Smart" method: \( E(S) = E(X_1) + E(X_2) = 2 \cdot 3.5 = 7 \)
(since \( E(X_i) = 3.5 \), from last time).

Generalization: I roll a die \( n \) times,

\[ X_i = \text{number shown i\textsuperscript{th} time} \]
\[ S = X_1 + X_2 + \ldots + X_n \]
\[ E(S) = n \cdot E(X_1) = n \cdot 3.5 \]

(Computing the distribution of \( S \) is very complicated).
• Example: $X_1, X_2, \ldots, X_n$ independent Bernoulli ($p$) random variables.

i.e. $X_i = \begin{cases} 
1 & \text{with probability } p \\
0 & \text{with probability } 1-p
\end{cases}$

so $E(X_i) = 0 \cdot (1-p) + 1 \cdot p = p$. Define:

$S = X_1 + X_2 + \ldots + X_n$, so the distribution of $S$
is $P(S=k) = \binom{n}{k} p^k (1-p)^{n-k}$, Binomial($n$, $p$).

Therefore $E(S) = E(X_1) + E(X_2) + \ldots + E(X_n) = np$. (**)

Remark: the above example shows how to compute the expectation of a random variable (the Binomial($n$, $p$) random variable) by interpreting as the sum of $n$ independent random variables (the Bernoulli($p$) r.v.'s) for which the expectation is known.

Note, however, that the expression (**') holds even in the case where $X_1, \ldots, X_n$ are NOT independent. In this case, $S = \sum_{i=1}^{n} X_i$ would have a different distribution (i.e. not necessarily Binomial); however, its mean (i.e., its expected value) would be the same.