Lecture Outline, 3-27-09 and 3-30-09. See pp. 436-444.
# RANDOM VARIABLE

## boats sold

<table>
<thead>
<tr>
<th>Number</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>0.2</td>
</tr>
<tr>
<td>4</td>
<td>0.3</td>
</tr>
<tr>
<td>5</td>
<td>0.1</td>
</tr>
<tr>
<td>6</td>
<td>0.1</td>
</tr>
<tr>
<td>7</td>
<td>0.05</td>
</tr>
<tr>
<td>8</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Total probability = 1

- $P(\text{fewer than 3.7}) = 0.4$
- $P(4 \text{ to } 7) = 0.55$

Chapter 16
A random variable is just a **numerical function** over the outcomes of a probability experiment.
Definition of \( E \ X \)
\[ E \ X = \text{sum of value times probability } \times p(x). \]

Key properties
\[ E(a \ X + b) = a \ E(X) + b \]
\[ E(X + Y) = E(X) + E(Y) \text{ (always, if such exist)} \]

a. \[ E(\text{sum of 13 dice}) = 13 \ E(\text{one die}) = 13(3.5). \]
b. \[ E(0.82 \text{ Ford US} + \text{Ford Germany} - 20M) \]
\[ = 0.82 \ E(\text{Ford US}) + E(\text{Ford Germany}) - 20M \]
regardless of any possible dependence.
## Total of 2 Dice

<table>
<thead>
<tr>
<th>Probability</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1/36</td>
</tr>
<tr>
<td>3</td>
<td>2/36</td>
</tr>
<tr>
<td>4</td>
<td>3/36</td>
</tr>
<tr>
<td>5</td>
<td>4/36</td>
</tr>
<tr>
<td>6</td>
<td>5/36</td>
</tr>
<tr>
<td>7</td>
<td>6/36</td>
</tr>
<tr>
<td>8</td>
<td>5/36</td>
</tr>
<tr>
<td>9</td>
<td>4/36</td>
</tr>
<tr>
<td>10</td>
<td>3/36</td>
</tr>
<tr>
<td>11</td>
<td>2/36</td>
</tr>
<tr>
<td>12</td>
<td>1/36</td>
</tr>
</tbody>
</table>

**Sum:**

\[
\text{sum} = 1 \\
\frac{252}{36} = 7
\]

The expected value of the total, \( E(\text{total}) \), is just twice the 3.5 avg for one die.
**boats/month**

<table>
<thead>
<tr>
<th>boats/month</th>
<th>probability</th>
<th>product</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>3</td>
<td>0.2</td>
<td>0.6</td>
</tr>
<tr>
<td>4</td>
<td>0.3</td>
<td>1.2</td>
</tr>
<tr>
<td>5</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>6</td>
<td>0.1</td>
<td>0.6</td>
</tr>
<tr>
<td>7</td>
<td>0.05</td>
<td>0.35</td>
</tr>
<tr>
<td>8</td>
<td>0.05</td>
<td>0.4</td>
</tr>
<tr>
<td><strong>total</strong></td>
<td><strong>1</strong></td>
<td><strong>4.05</strong></td>
</tr>
</tbody>
</table>

We average **4.05 boats per month**

\[ \text{E(number of boats this month)} \]
A random variable is just a **numerical function** over the outcomes of a probability experiment.
Expected return from policy “just drill” is the probability weighted average (NET) return
\[ E(\text{NET}) = (0.3) \times 270 + (0.7) \times (-130) = 81 - 91 = -10. \]

**just drill**

- \(0.3\) oil
- \(0.7\) no oil

Net return from policy “just drill.”
- \(-130 + 400 = 270\) drill oil
- \(-130 + 0 = -130\) drill no-oil

\[ E(X) = -10 \]
A test costing 20 is available. This test has:

\[ P(\text{test + | oil}) = 0.9 \]
\[ P(\text{test + | no-oil}) = 0.4. \]

Is it worth 20 to test first?
**EXPECTED RETURN IF WE "TEST FIRST"

<table>
<thead>
<tr>
<th>net return</th>
<th>prob</th>
<th>prod</th>
</tr>
</thead>
<tbody>
<tr>
<td>oil+</td>
<td>-20</td>
<td>130</td>
</tr>
<tr>
<td></td>
<td>+400</td>
<td></td>
</tr>
<tr>
<td></td>
<td>=250</td>
<td></td>
</tr>
<tr>
<td>oil-</td>
<td>-20</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>+0</td>
<td>=-20</td>
</tr>
<tr>
<td>no oil+</td>
<td>-20</td>
<td>130</td>
</tr>
<tr>
<td></td>
<td>+0</td>
<td>=-150</td>
</tr>
<tr>
<td>no oil-</td>
<td>-20</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>+0</td>
<td>=-20</td>
</tr>
</tbody>
</table>

**drill only if the test is +**

\[
\text{E(NET)} = .27 \times (250) - .03 \times (20) - .28 \times (150) - .42 \times (20)
\]

\[
= 16.5 \text{ (for the “test first” policy).}
\]

This average return is much preferred over the

\[
\text{E(NET)} = -10 \text{ of the “just drill” policy.}
\]
## Variance and s.d. of boats/month

<table>
<thead>
<tr>
<th>( x )</th>
<th>( p(x) )</th>
<th>( x ) ( p(x) )</th>
<th>( x^2 ) ( p(x) )</th>
<th>( (x-4.05)^2 ) ( p(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.2</td>
<td>0.4</td>
<td>0.8</td>
<td>0.8405</td>
</tr>
<tr>
<td>3</td>
<td>0.2</td>
<td>0.6</td>
<td>1.8</td>
<td>0.2205</td>
</tr>
<tr>
<td>4</td>
<td>0.3</td>
<td>1.2</td>
<td>4.8</td>
<td>0.0005</td>
</tr>
<tr>
<td>5</td>
<td>0.1</td>
<td>0.5</td>
<td>2.5</td>
<td>0.09025</td>
</tr>
<tr>
<td>6</td>
<td>0.1</td>
<td>0.6</td>
<td>3.6</td>
<td>0.38025</td>
</tr>
<tr>
<td>7</td>
<td>0.05</td>
<td>0.35</td>
<td>2.45</td>
<td>0.435125</td>
</tr>
<tr>
<td>8</td>
<td>0.05</td>
<td>0.4</td>
<td>3.2</td>
<td>0.780125</td>
</tr>
</tbody>
</table>

| total | 1.00 | 4.05 | 19.15 | 2.7475 |

### Terminology
- **E X** = mean
- **E \( X^2 \)** = mean of squares
- **E (X - E X)^2** = variance = mean of sq dev

\[
\text{s.d.} = \sqrt{2.7475} = \sqrt{19.15 - 4.05^2} = 1.6576
\]
**VARIANCE AND STANDARD DEVIATION**

\[
\text{Var}(X) = \text{def } E (X - E X)^2 = \text{comp } E (X^2) - (E X)^2
\]

i.e. \( \text{Var}(X) \) is the expected square deviation of r.v. \( X \) from its own expectation.

**Caution:** The computing formula (right above), although perfectly accurate mathematically, is sensitive to rounding errors.

**Key properties:**

\[
\text{Var}(a X + b) = a^2 \text{Var}(X) \text{ (b has no effect)}.
\]

\[
\text{sd}(a X + b) = |a| \text{sd}(X).
\]

\[
\text{VAR}(X + Y) = \text{Var}(X) + \text{VAR}(Y) \text{ if } X \text{ ind of } Y.
\]
Random variables $X, Y$ are INDEPENDENT if

$$p(x, y) = p(x) p(y) \text{ for all possible values } x, y.$$ 

If random variables $X, Y$ are INDEPENDENT

$$E (X Y) = (E X) (E Y) \text{ echoing the above.}$$

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$
Venture one returns random variable $X$ per $1$ investment. This $X$ is termed the “price relative.” This random $X$ may in turn be reinvested in venture two which returns random random variable $Y$ per $1$ investment. The return from $1$ invested at the outset is the product random variable $XY$.

If INDEPENDENT, $E( X Y ) = (E X)(E Y)$. 
EXAMPLE:

<table>
<thead>
<tr>
<th>$x$</th>
<th>$p(x)$</th>
<th>$x \cdot p(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>0.3</td>
<td>0.24</td>
</tr>
<tr>
<td>1.2</td>
<td>0.5</td>
<td>0.60</td>
</tr>
<tr>
<td>1.5</td>
<td>0.2</td>
<td>0.30</td>
</tr>
</tbody>
</table>

$E(X) = 1.14$

WE AVERAGE 14% PER PERIOD

BUT YOU WILL NOT EARN 14%. Simply put, the average is not a reliable guide to real returns in the case of exponential growth.
EXEMPLARY:  

\[
\begin{array}{ccc}
 x & p(x) & \log_e[x] \\ 
0.8 & 0.3 & -0.029073 \\
1.2 & 0.5 & 0.039591 \\
1.5 & 0.2 & 0.035218 \\
\end{array}
\]

\[
E \log_e[X] = 0.105311
\]

\[
e^{0.105311} \approx 1.11106
\]

With INDEPENDENT \underline{plays} your RANDOM return will compound at 11.1% not 14%.  

(more about this later in the course)
COMPARING $1.14^n$ WITH THREE RANDOM EVOLUTIONS

you can see that 14% exceeds reality
Poisson Distribution

Governing Counts of Rare Events

\[ p(x) = e^{-\text{mean}} \text{mean}^x / x! \]

for \( x = 0, 1, 2, \ldots \) ad infinitum
Poisson first best thing:

THE FIRST BEST THING ABOUT THE POISSON IS THAT THE MEAN ALONE TELLS US THE ENTIRE DISTRIBUTION!

note: Poisson sd = root(mean)
Poisson Cookies
400 raisins
144 COOKIES
mix well

\[ E X = \frac{400}{144} \sim 2.78 \] raisins per cookie
\[ sd = \text{root(mean)} = 1.67 \]
(for Poisson)
Poisson Cookies

e.g. $X =$ number of raisins in MY cookie. Batter has 400 raisins and makes 144 cookies.

$$E \ X = \frac{400}{144} \sim 2.78 \text{ per cookie}$$

$$p(x) = e^{-\text{mean}} \ \text{mean}^x / x!$$

e.g. $p(2) = e^{-2.78} \ 2.78^2 / 2! \sim 0.24$

(around 24% of cookies have 2 raisins)
Poisson
second best thing
THE SECOND BEST THING ABOUT
THE POISSON IS THAT FOR A MEAN
AS SMALL AS 3 THE NORMAL
APPROXIMATION WORKS WELL.

1.67 = sd = root(mean)
Special to Poisson

mean 2.78
Poison at cards

e.g. $X =$ number of times ace of spades turns up in 104 independent tries (i.e. from full deck)

$X \sim$ Poisson with mean 2

$p(x) = e^{-\text{mean}} \frac{\text{mean}^x}{x!}, \ x=0 \ldots$

$p(3) = e^{-2} \frac{2^3}{3!} \sim 0.182205$
Poisson in Risk

AVERAGE 127.8 ACCIDENTS PER MO.

E X = 127.8 accidents

If Poisson then sd = \( \sqrt{127.8} \) = 11.3049 and the approx dist is:

\[ \sim \]

sd = \( \sqrt{\text{mean}} \) = 11.3

Special to Poisson

mean 127.8 accidents
e.g. $X = \text{number of times ace of spades turns up}$
in 104 deals of 1 card from a shuffled full deck.

**Binomial** $(n=104, \ p = 1/52)$

$$p(x) = \binom{n}{x} \cdot p^x q^{n-x}, \ n = 0 \text{ to } n.$$  

$p(3) = \frac{(104!)/(3! \ 101!)}{(1/52)^3(51/52)^{49}} \approx 0.182205$

Agrees with Poisson approximation of binomial!
n = 10, p = 0.4
mean = n p = 4
sd = \sqrt{n p q} \sim 1.55
n = 30, p = 0.4
mean = n p = 12
sd = root(n p q)
~ 2.683
Normal Approx of Binomial

\[ n = 100, \ p = 0.4 \]
\[ \text{mean} = n \ p = 40 \]
\[ \text{sd} = \sqrt{n \ p \ q} \]
\[ \sim 4.89898 \]