Discrete & Continuous Distributions
Mixtures
& Expectations
The Kolmogorov Axioms of probability theory

1) \( \Pr(\Omega) = 1 \)

2) If \( V \) is a countable collection of disjoint events:
   \[ V = \{A_1, A_2, \ldots\}, \quad \forall i \neq j, \quad A_i \cap A_j = \emptyset \]

   Then Probability of the union is equal to the sum of the probabilities:
   \[ \Pr \left( \bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} \Pr(A_i) \]
The uniform distribution over \([0,1]\)

\[
0 \leq x_1, x_2, \ldots \leq 1 \\
P(\{x_1\}) = 0 \\
P(\{x_1, x_2\}) = 0 \\
P(\{x_1, x_2, x_3, \ldots\}) = 0 \\
\text{But, if } 0 \leq a < b \leq 1 \\
P([a, b]) = b - a
\]

Prob of other sets:

Construct from countable unions & intersections of intervals
\[ U(A, B) = \text{The Uniform distribution over the segment } [A, B] \]

\[ U(A, B) \text{ is defined by assigning probability } \]

to every segment \([a, b]\) where \(A \leq a \leq b \leq B\) \((A < B)\)

\[ \Pr([a, b]) = \Pr((a, b)) = \frac{b - a}{B - A} \]
Let's calculate the probability of some sets with respect to the uniform distribution

Fix the probability distribution $U(-1,1)$

$P([-1/3, 1/3]) = \frac{(1/3 - -1/3)}{(1 - -1)} = \frac{2/3}{2} = \frac{1}{3}$

$P([-1, 0]) = $

$P([-2, 0]) = $

$P([-3, 2]) = $

$P([0, 2]) = $

$P([-2, -1/2] U [1/2, 2]) = $
uniform density:

general density

\[ P([A,B]) = \int_{A}^{B} f(x) \, dx \]

\( \forall x \quad f(x) \geq 0 \quad (f(x) \text{ can be larger than 1}) \)

\[ \int_{-\infty}^{\infty} f(x) \, dx = 1 \]
PDF and CDF

The Probability Density function is shortened to PDF

Another popular representation of a distribution on the real is the CDF: Cumulative Distribution function

The CDF $F$ is defined as $F(a) \equiv \Pr(x \leq a)$

For density distributions, one can translate between PDF and CDF:

$$F(a) = \int_{-\infty}^{a} f(x) \, dx; \quad f(a) = \frac{dF(x)}{dx} \bigg|_{x=a}$$
Uniform Dist. $U(a, b)$
Endpoints: $a = 0$, $b = 1$
PDF:
\[
f(x) = \begin{cases} 
0 & x < a \\
\frac{1}{b-a} & a \leq x < b \\
0 & b \leq x
\end{cases}
\]

CDF:
\[
F(x) = \int_{-\infty}^{x} f(s)ds = \begin{cases} 
0 & x < a \\
\frac{x-a}{b-a} & a \leq x < b \\
1 & b \leq x
\end{cases}
\]
Standard Normal $N(0,1)$

PDF:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

CDF:

$$F(x) = \int_{-\infty}^{x} f(s)ds = 1 - Q(x)$$
**Shifted and Scaled Normal** \( N(\mu, \sigma) \)

**Shift:** \( \mu = 1 \)  **scale:** \( \sigma = 0.1 \)

**PDF:**

\[
f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}
\]

**CDF:**

\[
F(x) = \int_{-\infty}^{x} f(s) ds = 1 - Q\left(\frac{x-\mu}{\sigma}\right)
\]
**Exponential Distribution** \( \text{Exp}(a, \lambda) \)

*Shift: a  Scale: \( \lambda > 0 \)*

**PDF:**

\[
f(x) = \begin{cases} 
0, & x < a \\
\lambda e^{-\lambda(x-a)}, & x \geq a 
\end{cases}
\]

**CDF:**

\[
F(x) = \begin{cases} 
0, & x < a \\
1 - e^{-\lambda(x-a)}, & x \geq a 
\end{cases}
\]
Point-mass distribution $PM(a)$

Shift: $a$

PMF:

$P(a) = 1$

CDF:

$$F(x) = \int_{-\infty}^{x} f(s) ds = \begin{cases} 0 & x < a \\ 1 & x \geq a \end{cases}$$
(a) A wheel with four outcomes

(b) A wheel with infinitely many outcomes
A wheel with uncountably many outcomes

- **Point-mass distribution**: $P(b) = 1/4$, $P(a) = 3/4$
- **Density distribution**: uniform over $[a, b]$
- **Density distribution**: non-uniform
- **Mixture of point-mass and non-uniform density**
density distributions vs. Point-Mass distribution

Point mass distributions assign non-zero probability to individual points.
PM(a) ---- P(X=a)=1

Density distributions assign non-zero probability to segments.

The probability of any single point under a density distribution is zero.
=> as a result P([a,b])=P((a,b))=P([a,b))=P((a,b])

=> the probability of any countable set is zero.

=> for example the probability of all rational numbers in [0,1], under
the uniform distribution over [0,1] is zero!!!

In other words, if you pick a random number from U(0,1)
the probability that it is a rational number is zero !!!
Mixture Distributions
Mixtures distributions

$p_1 U(0,1) + p_2 PM(0) + p_3 U(-3,3) + p_4 PM(2)$

Choose which distribution

1. $U(0,1)$
2. $PM(0)$
3. $U(-3,3)$
4. $PM(2)$
\[ U((-1, +1)) \to 0.2 \to U(0, 4) \]

![Graphs showing distributions](image)
$U(-1,1)$

$Exp(1.5, 0.5)$

$0.3$

$0.7$
\[ .1U(0,1) + .2PM(0) + .3U(-3,3) + .4PM(2) \]

\[ F(-3) = 0; \quad F(-.01) \approx .5 \times .3 = .15 \]

\[ F(0) = .35; \quad F(1) = .35 + .1 + \frac{3}{6} = 0.5; \]

\[ F(1.99) \approx 0.5 + 0.05 = 0.55; \quad F(2) = 0.95 \]

\[ F(3) = 0.95 + \frac{3}{6} = 1.0 \]
Random Variables (RVs)
Sample space  = apples
An outcome is an apple
<table>
<thead>
<tr>
<th>Player</th>
<th>G</th>
<th>PA</th>
<th>AB</th>
<th>R</th>
<th>H</th>
<th>D</th>
<th>T</th>
<th>HR</th>
<th>RBI</th>
<th>BB</th>
<th>SO</th>
<th>BA</th>
<th>OBP</th>
<th>SLG</th>
<th>OPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mike Napoli</td>
<td>17</td>
<td>64</td>
<td>49</td>
<td>13</td>
<td>16</td>
<td>4</td>
<td>0</td>
<td>6</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>.327</td>
<td>.484</td>
<td>.776</td>
<td>1.260</td>
</tr>
<tr>
<td>Josh Donaldson</td>
<td>20</td>
<td>88</td>
<td>72</td>
<td>17</td>
<td>28</td>
<td>8</td>
<td>0</td>
<td>5</td>
<td>15</td>
<td>14</td>
<td>11</td>
<td>.389</td>
<td>.500</td>
<td>.708</td>
<td>1.209</td>
</tr>
<tr>
<td>Hunter Pence</td>
<td>21</td>
<td>92</td>
<td>78</td>
<td>19</td>
<td>25</td>
<td>4</td>
<td>0</td>
<td>9</td>
<td>26</td>
<td>13</td>
<td>14</td>
<td>.321</td>
<td>.413</td>
<td>.718</td>
<td>1.131</td>
</tr>
<tr>
<td>Matt Carpenter</td>
<td>21</td>
<td>102</td>
<td>88</td>
<td>23</td>
<td>35</td>
<td>11</td>
<td>2</td>
<td>1</td>
<td>11</td>
<td>12</td>
<td>19</td>
<td>.398</td>
<td>.480</td>
<td>.602</td>
<td>1.083</td>
</tr>
<tr>
<td>Ryan Zimmerman</td>
<td>21</td>
<td>97</td>
<td>90</td>
<td>22</td>
<td>29</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>16</td>
<td>7</td>
<td>8</td>
<td>.322</td>
<td>.371</td>
<td>.711</td>
<td>1.082</td>
</tr>
<tr>
<td>Freddi Freeman</td>
<td>20</td>
<td>85</td>
<td>73</td>
<td>15</td>
<td>26</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>17</td>
<td>10</td>
<td>15</td>
<td>.356</td>
<td>.435</td>
<td>.644</td>
<td>1.079</td>
</tr>
<tr>
<td>Michael Cuddyer</td>
<td>16</td>
<td>67</td>
<td>62</td>
<td>8</td>
<td>26</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>14</td>
<td>4</td>
<td>10</td>
<td>.419</td>
<td>.448</td>
<td>.629</td>
<td>1.077</td>
</tr>
<tr>
<td>Adam Lind</td>
<td>18</td>
<td>60</td>
<td>55</td>
<td>11</td>
<td>16</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>17</td>
<td>5</td>
<td>11</td>
<td>.291</td>
<td>.350</td>
<td>.709</td>
<td>1.050</td>
</tr>
<tr>
<td>Andrew McCutchen</td>
<td>20</td>
<td>83</td>
<td>66</td>
<td>13</td>
<td>22</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>13</td>
<td>11</td>
<td>.333</td>
<td>.470</td>
<td>.576</td>
<td>1.046</td>
</tr>
<tr>
<td>Prince Fielder</td>
<td>20</td>
<td>86</td>
<td>79</td>
<td>11</td>
<td>16</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>14</td>
<td>6</td>
<td>13</td>
<td>.380</td>
<td>.419</td>
<td>.620</td>
<td>1.039</td>
</tr>
<tr>
<td>Shin-Soo Choo</td>
<td>18</td>
<td>85</td>
<td>60</td>
<td>15</td>
<td>18</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>10</td>
<td>21</td>
<td>12</td>
<td>.300</td>
<td>.488</td>
<td>.550</td>
<td>1.038</td>
</tr>
<tr>
<td>Paul Goldschmidt</td>
<td>21</td>
<td>92</td>
<td>80</td>
<td>12</td>
<td>27</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>19</td>
<td>11</td>
<td>19</td>
<td>.338</td>
<td>.424</td>
<td>.613</td>
<td>1.036</td>
</tr>
<tr>
<td>Moises Sierra</td>
<td>19</td>
<td>67</td>
<td>63</td>
<td>8</td>
<td>22</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>4</td>
<td>14</td>
<td>.349</td>
<td>.388</td>
<td>.619</td>
<td>1.007</td>
</tr>
<tr>
<td>Jocjim Pinto</td>
<td>16</td>
<td>62</td>
<td>58</td>
<td>9</td>
<td>21</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>9</td>
<td>4</td>
<td>12</td>
<td>.362</td>
<td>.403</td>
<td>.603</td>
<td>1.007</td>
</tr>
<tr>
<td>Mike Trout</td>
<td>21</td>
<td>95</td>
<td>71</td>
<td>16</td>
<td>21</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>10</td>
<td>23</td>
<td>21</td>
<td>.296</td>
<td>.474</td>
<td>.521</td>
<td>.995</td>
</tr>
<tr>
<td>Yoenis Cespedes</td>
<td>19</td>
<td>80</td>
<td>77</td>
<td>12</td>
<td>26</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>19</td>
<td>2</td>
<td>19</td>
<td>.338</td>
<td>.363</td>
<td>.623</td>
<td>.986</td>
</tr>
<tr>
<td>Matt Holliday</td>
<td>20</td>
<td>92</td>
<td>76</td>
<td>14</td>
<td>28</td>
<td>6</td>
<td>0</td>
<td>2</td>
<td>20</td>
<td>14</td>
<td>13</td>
<td>.368</td>
<td>.457</td>
<td>.526</td>
<td>.983</td>
</tr>
<tr>
<td>David Ortiz</td>
<td>20</td>
<td>91</td>
<td>76</td>
<td>18</td>
<td>21</td>
<td>9</td>
<td>0</td>
<td>5</td>
<td>16</td>
<td>13</td>
<td>16</td>
<td>.276</td>
<td>.385</td>
<td>.592</td>
<td>.977</td>
</tr>
<tr>
<td>Chase Headley</td>
<td>17</td>
<td>67</td>
<td>56</td>
<td>8</td>
<td>15</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>9</td>
<td>10</td>
<td>12</td>
<td>.268</td>
<td>.388</td>
<td>.571</td>
<td>.959</td>
</tr>
<tr>
<td>Matt Adams</td>
<td>19</td>
<td>72</td>
<td>69</td>
<td>13</td>
<td>21</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>14</td>
<td>3</td>
<td>21</td>
<td>.304</td>
<td>.333</td>
<td>.623</td>
<td>.957</td>
</tr>
<tr>
<td>Joey Votto</td>
<td>20</td>
<td>94</td>
<td>73</td>
<td>12</td>
<td>22</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>9</td>
<td>20</td>
<td>17</td>
<td>.301</td>
<td>.447</td>
<td>.507</td>
<td>.954</td>
</tr>
<tr>
<td>Eric Hosmer</td>
<td>20</td>
<td>87</td>
<td>78</td>
<td>12</td>
<td>27</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>12</td>
<td>9</td>
<td>18</td>
<td>.346</td>
<td>.414</td>
<td>.526</td>
<td>.939</td>
</tr>
<tr>
<td>Wil Myers</td>
<td>20</td>
<td>84</td>
<td>78</td>
<td>10</td>
<td>24</td>
<td>9</td>
<td>0</td>
<td>4</td>
<td>10</td>
<td>6</td>
<td>18</td>
<td>.308</td>
<td>.357</td>
<td>.577</td>
<td>.934</td>
</tr>
<tr>
<td>Giancarlo Stanton</td>
<td>20</td>
<td>85</td>
<td>72</td>
<td>12</td>
<td>19</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>16</td>
<td>11</td>
<td>27</td>
<td>.264</td>
<td>.376</td>
<td>.556</td>
<td>.932</td>
</tr>
<tr>
<td>Desmond Jennings</td>
<td>21</td>
<td>83</td>
<td>68</td>
<td>7</td>
<td>19</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>13</td>
<td>13</td>
<td>16</td>
<td>.279</td>
<td>.398</td>
<td>.529</td>
<td>.927</td>
</tr>
</tbody>
</table>

**Outcome space:** all possible performances of baseball hitters for a month

**Outcome:** The performance of a particular player

**Random variables:** measures of performance: G, PA, AB ...

**Events:** More than 8 home runs, OPS higher than 1.0, 1.1, 1.2, ...
Event & RVs

- From RV to event
  \[ A = \{ \omega \in \Omega \mid X(\omega) > 5 \} \]

- From event to RV
  \[ X = \begin{cases} 
  1 & \text{if } \omega \in A \\
  0 & \text{if } \omega \notin A 
\end{cases} \]

RVs \( X(\omega), Y(\omega) \) are independent if

\[ \forall A, B, A \text{ defined using } X \\
B \text{ defined using } Y \\
A, B \text{ are independent} \]
Joint distribution of two independent random variables

<table>
<thead>
<tr>
<th></th>
<th>X=1</th>
<th>X=2</th>
<th>X=10</th>
<th>P(Y=y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y=−1</td>
<td>1/12</td>
<td>1/12</td>
<td>2/12</td>
<td>4/12=1/3</td>
</tr>
<tr>
<td>Y=+1</td>
<td>2/12</td>
<td>2/12</td>
<td>4/12</td>
<td>8/12=2/3</td>
</tr>
<tr>
<td>P(X=x)</td>
<td>3/12=1/4</td>
<td>3/12=1/4</td>
<td>6/12=1/2</td>
<td></td>
</tr>
</tbody>
</table>
Joint distribution of two dependent random variables

<table>
<thead>
<tr>
<th></th>
<th>X=1</th>
<th>X=2</th>
<th>X=10</th>
<th>P(Y=y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y=-1</td>
<td>1/12</td>
<td>2/12</td>
<td>1/12</td>
<td>4/12=1/3</td>
</tr>
<tr>
<td>Y=+1</td>
<td>2/12</td>
<td>1/12</td>
<td>5/12</td>
<td>8/12=2/3</td>
</tr>
<tr>
<td>P(X=x)</td>
<td>3/12=1/4</td>
<td>3/12=1/4</td>
<td>6/12=1/2</td>
<td></td>
</tr>
</tbody>
</table>

Marginals
Expected Value
Figure 2.22: A weight system representing the probability distribution for $X$. The string holds the distribution at the mean to keep the system balanced.

<table>
<thead>
<tr>
<th>$i$</th>
<th>$x_i$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X = x_i)$</td>
<td>0.20</td>
<td>0.55</td>
<td>0.25</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

$E(X) = 0 \times P(X = 0) + 137 \times P(X = 137) + 170 \times P(X = 170)$

$= 0 \times 0.20 + 137 \times 0.55 + 170 \times 0.25 = 117.85$

**Expected value of a Discrete Random Variable**

If $X$ takes outcomes $x_1, \ldots, x_k$ with probabilities $P(X = x_1), \ldots, P(X = x_k)$, the expected value of $X$ is the sum of each outcome multiplied by its corresponding probability:

\[
E(X) = x_1 \times P(X = x_1) + \cdots + x_k \times P(X = x_k)
\]

\[
= \sum_{i=1}^{k} x_i P(X = x_i)
\]

(2.71)

The Greek letter $\mu$ may be used in place of the notation $E(X)$. 
Figure 2.23: A continuous distribution can also be balanced at its mean.

\[ E(x) = \int_{-\infty}^{\infty} s f(s) \, ds \]
Expected Value

- Suppose $X$ is a discrete random variable $P(X = a_i) = p_i$
  - The expected value of $X$ is $E(X) = \sum_{i=1}^{n} p_i a_i$
- Suppose $X$ is a continuous random variable with density $f$
  - The expected value of $X$ is $E(X) = \int_{-\infty}^{+\infty} f(x) dx$
- $E(X)$ is a property of the distribution, it is not a random variable.
- The average is a random variable:
  - $Average(x_1, x_2, ..., x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i$
- When $n$ is large, the average tends to be close to the mean.
Example - Binary random variables:
Let $X_1, X_2, \ldots, X_{100}$

Be independent binary random variables: $P(X_i = 0) = P(X_i = 1) = \frac{1}{2}$

Let $S = \frac{1}{100} \sum_{i=1}^{100} X_i$, $S$ is the ________, $S$ is/is-not a random variable?

$E(X_i) = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2}$, $E(X_i)$ is/is-not a random variable?

What is $E(S)$?
Rules for expected value:

1. If $a, b$ are constants and $X$ is a random variable then
   \[ E(aX + b) = aE(X) + b \]

2. If $X, Y$ are random variables (dependent or independent)
   \[ E(X + Y) = E(X) + E(Y) \]
   \[ \text{—> what is } E(aX + bY + c) = ? \]

3. If the distribution of the RV $X$ is a mixture of two distributions:
   \[ P = pP_1 + (1 - p)P_2 \]
   then
   \[ E_P(X) = pE_{P_1}(X) + (1 - p)E_{P_2}(X) \]

So now, $S = \frac{1}{100} \sum_{i=1}^{100} X_i$, what is $E(S)$?
next Class
Expectation & Variance
CDFs Vs Histograms