Continuous random variables

*Discrete* random variable: takes values in a finite or countable set, e.g.

\[ X \in \{1, 2, ..., 6\} \] with equal probability

\[ X \] is positive integer \( i \) with probability \( 2^{-i} \)

*Continuous* random variable: takes values in an uncountable set, e.g.

\[ X \] is the weight of a random person (a real number)

\[ X \] is a randomly selected point inside a unit square

\[ X \] is the waiting time until the next packet arrives at the server
f(x) : the \textit{probability density function} (or simply “density”)

\[ F(a) = \int_{-\infty}^{a} f(x) \, dx \]

\[ P(X \leq a) = F(a) \text{: the \textit{cumulative distribution function} (or simply “distribution”)} \]

\[ P(a < X \leq b) = F(b) - F(a) \]

Need \[ f(x) \geq 0, \quad \int_{-\infty}^{+\infty} f(x) \, dx \quad (= F(+\infty)) \quad = 1 \]

A key relationship:

\[ f(x) = \frac{d}{dx} F(x), \text{ since } F(a) = \int_{-\infty}^{a} f(x) \, dx, \]
Densities are not probabilities; e.g. may be > 1

\[ P(x = a) = P(a \leq X \leq a) = F(a) - F(a) = 0 \]

I.e., the probability that a continuous random variable falls at a specified point is zero.

\[ P(a - \varepsilon/2 \leq X \leq a + \varepsilon/2) = F(a + \varepsilon/2) - F(a - \varepsilon/2) \approx \varepsilon \cdot f(a) \]

I.e., The probability that it falls near that point is proportional to the density; in a large random sample, expect more samples where density is higher (hence the name “density”).
Much of what we did with discrete r.v.s carries over almost unchanged, with $\sum_x \ldots$ replaced by $\int \ldots \, dx$

E.g.

For discrete r.v. $X$, $E[X] = \sum_x x p(x)$

For continuous r.v. $X$, $E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \, dx$

Why?

(a) We define it that way

(b) The probability that $X$ falls “near” $x$, say within $x \pm dx/2$, is $\approx f(x) \, dx$, so the “average” $X$ should be $\approx \sum x f(x) \, dx$ (summed over grid points spaced $dx$ apart on the real line) and the limit of that as $dx \to 0$ is $\int xf(x) \, dx$
Let \( f(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases} \)

\[
F(a) = \int_{-\infty}^{a} f(x) \, dx
= \begin{cases} 0 & \text{if } a \leq 0 \\ a & \text{if } 0 < a \leq 1 \quad (\text{since } a = \int_{0}^{a} 1 \, dx) \\ 1 & \text{if } 1 < a \end{cases}
\]

\[
E[X] = \int_{-\infty}^{\infty} x f(x) \, dx = \int_{0}^{1} x \, dx = \frac{x^2}{2} \bigg|_{0}^{1} = \frac{1}{2}
\]

\[
E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) \, dx = \int_{0}^{1} x^2 \, dx = \frac{x^3}{3} \bigg|_{0}^{1} = \frac{1}{3}
\]

\[
\text{Var}[X] = E[X^2] - (E[X])^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \quad (\sigma \approx 0.29)
\]
properties of expectation

Linearity

\[ E[aX + b] = aE[X] + b \]

\[ E[X + Y] = E[X] + E[Y] \]

Functions of a random variable

\[ E[g(X)] = \int g(x)f(x)\,dx \]

Alternatively, let \( Y = g(X) \), find the density of \( Y \), say \( f_Y \), (see B&T 4.1; somewhat like r.v. slides 33-35) and directly compute \( E[Y] = \int yf_Y(y)\,dy \).
variance

Definition is same as in the discrete case

\[ \text{Var}[X] = \text{E}[(X-\mu)^2] \text{ where } \mu = \text{E}[X] \]

Identity still holds:

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

proof “same”
Let $f(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$

$$F(a) = \int_{-\infty}^{a} f(x) \, dx$$

= \begin{cases} 0 & \text{if } a \leq 0 \\ a & \text{if } 0 < a \leq 1 \quad (\text{since } a = \int_{0}^{a} 1 \, dx) \\ 1 & \text{if } 1 < a \end{cases}$$

$$E[X] = \int_{-\infty}^{\infty} x f(x) \, dx = \int_{0}^{1} x \, dx = \frac{x^2}{2} \bigg|_{0}^{1} = \frac{1}{2}$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) \, dx = \int_{0}^{1} x^2 \, dx = \frac{x^3}{3} \bigg|_{0}^{1} = \frac{1}{3}$$

$$\text{Var}[X] = E[X^2] - (E[X])^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \quad (\sigma \approx 0.29)$$
Continuous random variable $X$ has density $f(x)$, and

$$\Pr(a \leq X \leq b) = \int_a^b f(x) \, dx$$

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \, dx$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x) \, dx$$
$X \sim \text{Uni}(\alpha, \beta)$ is uniform in $[\alpha, \beta]$.

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & x \in [\alpha, \beta] \\ 0 & \text{otherwise} \end{cases}$$

The Uniform Density Function $\text{Uni}(0.5, 1.0)$
\( X \sim \text{Uni}(\alpha, \beta) \) is uniform in \([\alpha, \beta]\)

\[
f(x) = \begin{cases} 
\frac{1}{\beta - \alpha} & x \in [\alpha, \beta] \\
0 & \text{otherwise}
\end{cases}
\]

\[
\Pr(a \leq X \leq b) = \int_{a}^{b} f(x) \, dx = \frac{b - a}{\beta - \alpha}
\]

if \( \alpha \leq a \leq b \leq \beta \):

\[
E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \, dx = \frac{\alpha + \beta}{2}
\]
uniform random variable: example

\[ X \sim \text{Uni}(\alpha, \beta) \] is uniform in \([\alpha, \beta]\)

\[
f(x) = \begin{cases} 
\frac{1}{\beta - \alpha} & \text{if } x \in [\alpha, \beta] \\
0 & \text{otherwise}
\end{cases}
\]

You want to read a disk sector from a 7200rpm disk drive. Let \( T \) be the time you wait, in milliseconds, after the disk head is positioned over the correct track, until the desired sector rotates under the head.

\[ T \sim \text{Uni}(0, 8.33) \]
X ~ Exp(λ)

The Exponential Density Function

\[ f(x) = \begin{cases} 
\lambda e^{-\lambda x} & x \geq 0 \\
0 & x < 0 
\end{cases} \]
$X \sim \text{Exp}(\lambda)$

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

\[ E[X] = \frac{1}{\lambda} \quad \text{Var}[X] = \frac{1}{\lambda^2} \]

\[ \Pr(X \geq t) = e^{-\lambda t} = 1 - F(t) \]

Memorylessness:

\[ \Pr(X > s + t \mid X > s) = \Pr(X > t) \]
Examples

Radioactive decay: How long until the next alpha particle?

Customers: how long until the next customer/packet arrives at the checkout stand/server?

Buses: How long until the next #71 bus arrives on the Ave?
   Yes, they have a schedule, but given the vagaries of traffic, riders with-bikes-and-baby-carriages, etc., can they stick to it?

Gambler’s fallacy: “I’m due for a win”

Relation to the Poisson: same process, different measures:
   Poisson: how many events in a fixed time;
   Exponential: how long until the next event

Relation to geometric: Geometric is discrete analog:
   How long to a Head, 1 flip per sec, prob p vs
   How long to a Head, 2 flips per sec, prob p/2, ...
   Limit is exponential with parameter p
X is a normal (aka Gaussian) random variable \( X \sim \mathcal{N}(\mu, \sigma^2) \)

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

\[
E[X] = \mu \quad \text{Var}[X] = \sigma^2
\]

The Standard Normal Density Function
Changing $\mu, \sigma$ 

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

Density at $\mu$ is $\approx 0.399/\sigma$
normal random variable

\(X\) is a normal random variable \(X \sim N(\mu, \sigma^2)\)

\[Y = aX + b\]

\[E[Y] = E[aX+b] = a\mu + b\]

\[\text{Var}[Y] = \text{Var}[aX+b] = a^2\sigma^2\]

\(Y \sim N(a\mu + b, a^2\sigma^2)\)

Important special case: \(Z = (X-\mu)/\sigma \sim N(0,1)\)

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

\(Z \sim N(0,1)\) “standard (or unit) normal”

Use \(\Phi(z)\) to denote CDF, i.e.

\[\Phi(z) = \Pr(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx\]

no closed form 😞
<table>
<thead>
<tr>
<th>z</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.06</th>
<th>0.07</th>
<th>0.08</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.5000</td>
<td>0.5040</td>
<td>0.5080</td>
<td>0.5120</td>
<td>0.5160</td>
<td>0.5199</td>
<td>0.5239</td>
<td>0.5279</td>
<td>0.5319</td>
</tr>
<tr>
<td>0.1</td>
<td>0.5398</td>
<td>0.5438</td>
<td>0.5478</td>
<td>0.5517</td>
<td>0.5557</td>
<td>0.5596</td>
<td>0.5636</td>
<td>0.5675</td>
<td>0.5714</td>
</tr>
<tr>
<td>0.2</td>
<td>0.5793</td>
<td>0.5832</td>
<td>0.5871</td>
<td>0.5910</td>
<td>0.5948</td>
<td>0.5987</td>
<td>0.6026</td>
<td>0.6064</td>
<td>0.6103</td>
</tr>
<tr>
<td>0.3</td>
<td>0.6179</td>
<td>0.6217</td>
<td>0.6255</td>
<td>0.6293</td>
<td>0.6331</td>
<td>0.6368</td>
<td>0.6406</td>
<td>0.6443</td>
<td>0.6480</td>
</tr>
<tr>
<td>0.4</td>
<td>0.6554</td>
<td>0.6593</td>
<td>0.6628</td>
<td>0.6664</td>
<td>0.6700</td>
<td>0.6736</td>
<td>0.6772</td>
<td>0.6808</td>
<td>0.6844</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6915</td>
<td>0.6950</td>
<td>0.6985</td>
<td>0.7019</td>
<td>0.7054</td>
<td>0.7088</td>
<td>0.7122</td>
<td>0.7157</td>
<td>0.7190</td>
</tr>
<tr>
<td>0.6</td>
<td>0.7257</td>
<td>0.7291</td>
<td>0.7324</td>
<td>0.7357</td>
<td>0.7389</td>
<td>0.7422</td>
<td>0.7455</td>
<td>0.7488</td>
<td>0.7517</td>
</tr>
<tr>
<td>0.7</td>
<td>0.7580</td>
<td>0.7611</td>
<td>0.7642</td>
<td>0.7673</td>
<td>0.7704</td>
<td>0.7734</td>
<td>0.7764</td>
<td>0.7793</td>
<td>0.7823</td>
</tr>
<tr>
<td>0.8</td>
<td>0.7881</td>
<td>0.7910</td>
<td>0.7939</td>
<td>0.7967</td>
<td>0.7995</td>
<td>0.8023</td>
<td>0.8051</td>
<td>0.8078</td>
<td>0.8106</td>
</tr>
<tr>
<td>0.9</td>
<td>0.8159</td>
<td>0.8186</td>
<td>0.8212</td>
<td>0.8238</td>
<td>0.8264</td>
<td>0.8289</td>
<td>0.8315</td>
<td>0.8340</td>
<td>0.8365</td>
</tr>
</tbody>
</table>

The Standard Normal Density Function

E.g., see B&T p155, p531
If $Z \sim N(\mu, \sigma)$ what is $P( \mu - \sigma < Z < \mu + \sigma )$?

$P( \mu - \sigma < Z < \mu + \sigma ) = \Phi(1) - \Phi(-1) \approx 68\%$

$P( \mu - 2\sigma < Z < \mu + 2\sigma ) = \Phi(2) - \Phi(-2) \approx 95\%$

$P( \mu - 3\sigma < Z < \mu + 3\sigma ) = \Phi(3) - \Phi(-3) \approx 99\%$
normal approximation to binomial

\[ X \sim \text{Bin}(n,p) \]
\[ E[X] = np \quad \text{Var}[X] = np(1-p) \]

Poisson approx: good for \( n \) large, \( p \) small (\( np \) constant)
Normal approx: For large \( n \), (\( p \) stays fixed):
\[ X \approx Y \sim \text{N}(E[X], \text{Var}[X]) = \text{N}(np, np(1-p)) \]
Normal approximation good when \( np(1-p) \geq 10 \)

**DeMoivre-Laplace Theorem:**
Let \( S_n \) = number of successes in \( n \) trials (with prob. \( p \)).
Then, as \( n \to \infty \):

\[
Pr \left( a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b \right) \to \Phi(b) - \Phi(a)
\]
normal approximation to binomial

- Normal(np, np(1-p))
- Binomial(n, p)
- Poisson(np)

n = 100
p = 0.5
normal approximation to binomial

Fair coin flipped 40 times. Probability of 20 heads?

Exact answer:
\[ P(X = 20) = \binom{40}{20} \left( \frac{1}{2} \right)^{40} \approx 0.1254 \]

Normal approximation:
\[ P(X = 20) = P(19.5 \leq X < 20.5) \]
\[ = P \left( \frac{19.5 - 20}{\sqrt{10}} \leq \frac{X - 20}{\sqrt{10}} < \frac{20.5 - 20}{\sqrt{10}} \right) \]
\[ \approx P \left( -0.16 \leq \frac{X - 20}{\sqrt{10}} < 0.16 \right) \]
\[ \approx \Phi(0.16) - \Phi(-0.16) \approx 0.1272 \]
the central limit theorem (CLT)

Consider i.i.d. (independent, identically distributed) random vars $X_1, X_2, X_3, \ldots$

$X_i$ has $\mu = E[X_i]$ and $\sigma^2 = \text{Var}[X_i]$

As $n \to \infty$,

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma \sqrt{n}} \to N(0, 1)$$

Restated: As $n \to \infty$,

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \sim N \left( \mu, \frac{\sigma^2}{n} \right)$$
How tall are you? Why?

Willie Shoemaker & Wilt Chamberlain
in the real world...

Human height is approximately normal.

Why might that be true?

R.A. Fisher (1918) noted it would follow from CLT if height were the sum of many independent random effects, e.g. many genetic factors (plus some environmental ones like diet). I.e., suggested part of mechanism by looking at shape of the curve. (WAY before anyone really knew what genes were...)
### Table 1. Sixty-Four Loci Showing Significant Evidence for Association with Adult Height, Identified with the Use of the IBC Array

<table>
<thead>
<tr>
<th>Locus Rank</th>
<th>Chr.</th>
<th>Candidate Gene*</th>
<th>SNP*</th>
<th>Effect Allele</th>
<th>MAF</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7q22</td>
<td>CDK6</td>
<td>rs4272</td>
<td>A</td>
<td>0.21</td>
<td>−0.46</td>
</tr>
<tr>
<td>2</td>
<td>6p21</td>
<td>HMGAI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>12q16</td>
<td>HMGAI2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>20q11</td>
<td>MMP24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>17q23</td>
<td>MAP3K3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>17q24</td>
<td>GH1-GH2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1p36</td>
<td>MEAP2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>15q26</td>
<td>IGFI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>7p22</td>
<td>GNA12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>17q23</td>
<td>TBX2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>12q22</td>
<td>Socs2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>9q22</td>
<td>PTCH1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>14q11</td>
<td>NFATC4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>15q26</td>
<td>ANCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>2q24</td>
<td>NPPC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>6p21</td>
<td>PPARD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>20q11</td>
<td>MYH7B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>19q13</td>
<td>IL11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 1. Continued

<table>
<thead>
<tr>
<th>Locus Rank</th>
<th>Chr.</th>
<th>Candidate Gene*</th>
<th>SNP*</th>
<th>Effect Allele</th>
<th>MAF</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>2p23</td>
<td>GCKR</td>
<td>rs780094</td>
<td>A</td>
<td>0.16</td>
<td>0.23</td>
</tr>
<tr>
<td>29</td>
<td>1q41</td>
<td>TGFβ2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>20q11</td>
<td>CDK5RAP1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>2p12</td>
<td>EIF2AK3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>19p13</td>
<td>INSR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>6q25</td>
<td>ESR1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>2q37</td>
<td>DIS2L</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>2q35</td>
<td>PLCD4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>1p36</td>
<td>RS56AK1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>15q21</td>
<td>CYP19A1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>9q31</td>
<td>SLC22A5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>7q15</td>
<td>JAZF1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>17p13</td>
<td>PDL2A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>1p22</td>
<td>PKN2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>7q22</td>
<td>CNOT4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### European Ancestry Phase I (up to 53,394)

- $P < 1 \times 10^{-10}$
- $P < 5 \times 10^{-8}$
- $P < 5 \times 10^{-7}$
- $P < 2 \times 10^{-6}$

<table>
<thead>
<tr>
<th>Locus Rank</th>
<th>Chr.</th>
<th>Candidate Gene*</th>
<th>SNP*</th>
<th>Effect Allele</th>
<th>MAF</th>
<th>Effect</th>
<th>p</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>54</td>
<td>1p22</td>
<td>COL24A1</td>
<td>rs2046159</td>
<td>A</td>
<td>0.16</td>
<td>0.23</td>
<td>$3.8 \times 10^{-5}$</td>
<td>2</td>
</tr>
<tr>
<td>55</td>
<td>1q23</td>
<td>DSG2</td>
<td>rs1129923</td>
<td>A</td>
<td>0.10</td>
<td>−0.25</td>
<td>$2.7 \times 10^{-4}$</td>
<td>0</td>
</tr>
<tr>
<td>56</td>
<td>10q22</td>
<td>MAT2A</td>
<td>rs7087278</td>
<td>A</td>
<td>0.18</td>
<td>0.22</td>
<td>$2.2 \times 10^{-4}$</td>
<td>0</td>
</tr>
<tr>
<td>57</td>
<td>2p15</td>
<td>PIP5R1</td>
<td>rs1822269</td>
<td>T</td>
<td>0.41</td>
<td>−0.14</td>
<td>$7.8 \times 10^{-4}$</td>
<td>9</td>
</tr>
<tr>
<td>58</td>
<td>7q36</td>
<td>ATG9B</td>
<td>rs1800783</td>
<td>A</td>
<td>0.38</td>
<td>−0.16</td>
<td>$2.0 \times 10^{-4}$</td>
<td>0</td>
</tr>
<tr>
<td>59</td>
<td>14q11</td>
<td>RCCL2</td>
<td>rs3210043</td>
<td>A</td>
<td>0.16</td>
<td>0.25</td>
<td>$9.7 \times 10^{-5}$</td>
<td>0</td>
</tr>
<tr>
<td>60</td>
<td>4p14</td>
<td>RFC1</td>
<td>rs11096991</td>
<td>T</td>
<td>0.35</td>
<td>0.15</td>
<td>$3.6 \times 10^{-4}$</td>
<td>0</td>
</tr>
</tbody>
</table>
in the real world...
in the real world...
in the real world...

Histogram of Daily Trading-Related Revenue* — Twelve Months Ended December 31, 2007

*Excludes daily profits and losses in the ABS CDO market, including recent subprime-related losses.
in the real world…
pdf and cdf

\[ f(x) = \frac{d}{dx} F(x) \quad F(a) = \int_{-\infty}^{a} f(x) \, dx \]

sums become integrals, e.g.

\[ E[X] = \sum_x x \, p(x) \quad E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \, dx \]

most familiar properties still hold, e.g.

\[ E[aX+bY+c] = aE[X]+bE[Y]+c \]

\[ \text{Var}[X] = E[X^2] - (E[X])^2 \]
continuous r.v.’s: summary

Three important examples

\[ X \sim \text{ Uni}(\alpha, \beta) \text{ uniform in } [\alpha, \beta] \]

\[ f(x) = \begin{cases} \frac{1}{\beta - \alpha} & x \in [\alpha, \beta] \\ 0 & \text{otherwise} \end{cases} \]

\[ E[X] = (\alpha + \beta)/2 \]
\[ \text{Var}[X] = (\alpha - \beta)^2/12 \]

\[ X \sim \text{ Exp}(\lambda) \text{ exponential} \]

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

\[ E[X] = \frac{1}{\lambda} \]
\[ \text{Var}[X] = \frac{1}{\lambda^2} \]

\[ X \sim \text{ N}(\mu, \sigma^2) \text{ normal (aka Gaussian)} \]

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

\[ E[X] = \mu \]
\[ \text{Var}[X] = \sigma^2 \]