4.2 Continuous Models

Horseshoe Crab (*Limulus polyphemus*)

- Not true crabs, but closely related to spiders and scorpions.
- “Living fossils” – existed since Carboniferous Period, ≈ 350 mya.
- Found primarily on Atlantic coast, with the highest concentration in Delaware Bay, where males and the much larger females congregate in large numbers on the beaches for mating, and subsequent egg-laying.

- **Pharmaceutical (and many other scientific) contributions!**
  Blue hemolymph (due to copper-based hemocyanin molecule) contains *amebocytes*, which produce a clotting agent that reacts with endotoxins found in the outer membrane of Gram-negative bacteria. Several East Coast companies have developed the *Limulus Amebocyte Lysate* (LAL) assay, used to detect bacterial contamination of drugs and medical implant devices, etc. Equal amounts of LAL reagent and test solution are mixed together, incubated at 37°C for one hour, then checked to see if gelling has occurred. Simple, fast, cheap, sensitive, uses very small amounts, and does not harm the animals… *probably.* (Currently, a moratorium exists on their harvesting, while population studies are ongoing…)

Photo courtesy of Bill Hall, bhall@udel.edu. Used with permission.
**Continuous Random Variable:**

\( X = \text{“Length (inches) of adult horseshoe crabs”} \)

**Sample 1**

\( n = 25; \) lengths measured to nearest inch

e.g., 10 in \([12, 16)\), 6 in \([16, 20)\), 9 in \([20, 24)\)

**Sample 2**

\( n = 1000; \) lengths measured to nearest \(\frac{1}{2}\) inch

e.g., 180 in \([12, 14)\), 240 in \([14, 16)\), etc.

**Examples:**

\[ P(16 \leq X < 20) = 0.24 \]

\[ P(16 \leq X < 20) = 0.16 + 0.12 = 0.28 \]

In the limit as \( n \to \infty \), the population distribution of \( X \) can be characterized by a continuous density curve, and formally described by a density function \( f(x) \geq 0 \).

Thus, \( P(a \leq X < b) = \int_a^b f(x) \, dx = \text{area under the density curve from } a \text{ to } b. \)
Definition: \( f(x) \) is a **probability density function** (pdf) for the continuous random variable \( X \) if, for all \( x \),

\[
f(x) \geq 0 \quad \text{AND} \quad \int_{-\infty}^{\infty} f(x) \, dx = 1.
\]

The **cumulative distribution function** (cdf) is defined as, for all \( x \),

\[
F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t) \, dt.
\]

Therefore, \( F \) increases monotonically and continuously from 0 to 1.

Furthermore, \( P(a \leq X \leq b) = \int_{a}^{b} f(x) \, dx = F(b) - F(a) \). \( \text{FTC}!! \)

The cumulative probability that \( X \) is less than or equal to some value \( x \) – i.e., \( P(X \leq x) \) – is characterized by:

1. the area under the graph of \( f \) up to \( x \), or
2. the height of the graph of \( F \) at \( x \).

**But note:** \( f(x) \) NO LONGER corresponds to the probability \( P(X = x) \) [which = 0, since \( X \) is here continuous], as it does for discrete \( X \).
Example 1: Uniform density

This is the trivial “constant function” over some fixed interval \([a, b]\). That is,  
\[
    f(x) = \frac{1}{b-a} \quad \text{for } a \leq x \leq b \quad \text{and } f(x) = 0 \text{ otherwise}.
\]
Clearly, the two criteria for being a valid density function are met: it is non-negative, and the (rectangular) area under its graph is equal to its base \((b - a) \times \text{height} \ (1 / b - a)\), which is indeed 1. Moreover, for any value of \(x\) in the interval \([a, b]\), the (rectangular) area under the graph up to \(x\) is equal to its base \((x - a) \times \text{height} \ (1 / b - a)\). That is, the cumulative distribution function (cdf) is given by  
\[
    F(x) = \frac{x-a}{b-a},
\]
the graph of which is a straight line connecting the left endpoint \((a, 0)\) to the right endpoint \((b, 1)\).

[[Note: Since \(f(x) = 0\) outside the interval \([a, b]\), the area beneath it contributes nothing to \(F(x)\) there; hence \(F(x) = 0\) if \(x < a\), and \(F(x) = 1\) if \(x > b\). Observe that, indeed, \(F\) increases \textit{monotonically} and \textit{continuously} from 0 to 1; the graphs show \(f(x)\) and \(F(x)\) over the interval \([1, 6]\), i.e., \(a = 1, b = 6\). \textit{Compare this example with the discrete version in section 3.1.}]]

Thus, for example, the probability \(P(2.6 \leq X \leq 3.8)\) is equal to the (rectangular) area under \(f(x)\) over that interval, or in terms of \(F(x)\), simply equal to the difference between the heights \(F(3.8) - F(2.6)\)  
\[
    = \frac{3.8-1}{5} - \frac{2.6-1}{5} = 0.56 - 0.32 = 0.24.
\]
Example 2: **Power density**  (A special case of the **Beta density**: $\beta = 1$)

For any fixed $p > 0$, let $f(x) = px^{p-1}$ for $0 < x < 1$. (Else, $f(x) = 0$.) This is a valid density function, since $f(x) \geq 0$ and

$$\int_{-\infty}^{\infty} f(x) \, dx = \int_{0}^{1} px^{p-1} \, dx = \left[ x^p \right]_{0}^{1} = 1 \checkmark.$$ 

The corresponding cdf is therefore $F(x) = \int_{-\infty}^{x} f(t) \, dt = \int_{0}^{x} p t^{p-1} \, dt = \left[ t^p \right]_{0}^{x} = x^p$

on $[0, 1]$. (And, as above, $F(x) = 0$ if $x < 0$, and $F(x) = 1$ if $x > 1$.) Again observe that $F$ indeed increases *monotonically* and *continuously* from 0 to 1, regardless of $f$; see graphs for $p = \frac{1}{2}, \frac{3}{2}, 3$.  (Note: $p = 1$ corresponds to the **uniform density** on $[0, 1]$.)
Example 3: Cauchy density

The function \( f(x) = \frac{1}{\pi \left( 1 + x^2 \right)} \) for \(-\infty < x < +\infty\) is a legitimate density function, since it satisfies the two criteria above: \( f(x) \geq 0 \) AND \( \int_{-\infty}^{\infty} f(x) \, dx = 1 \). (Verify it!) The cdf is therefore \( F(x) = \int_{-\infty}^{x} f(t) \, dt = \int_{-\infty}^{x} \frac{1}{\pi \left( 1 + t^2 \right)} \, dt = \frac{1}{\pi} \arctan \frac{1}{2} \) for \(-\infty < x < +\infty\).

Thus, for instance, \( P(0 \leq X \leq 1) = F(1) - F(0) = \left[ \frac{1}{\pi} \right. \frac{\pi}{4} + \left. \frac{1}{2} \right] - \left[ \frac{1}{\pi} \right. 0 + \left. \frac{1}{2} \right] = \frac{1}{4}.

Example 4: Exponential density

For any \( a > 0 \) fixed, \( f(x) = a e^{-ax} \) for \( x \geq 0 \) (and = 0 for \( x < 0 \)) is a valid density function, since it satisfies the two criteria. (Details are left as an exercise.) The corresponding cdf is given by \( F(x) = \int_{-\infty}^{x} f(t) \, dt = \int_{-\infty}^{x} a e^{-at} \, dt = 1 - e^{-ax} \), for \( x \geq 0 \) (and = 0 otherwise).

The case \( a = 1 \) is shown below.

Thus, for instance, \( P(X \leq 2) = F(2) = 1 - e^{-2} = 0.8647\), and \( P(0.5 \leq X \leq 2) = F(2) - F(0.5) = (1 - e^{-2}) - (1 - e^{-0.5}) = 0.8647 - 0.3935 = 0.4712\).
Exercise: (Another special case of the **Beta density**.) Sketch the graph of \( f(x) = 6x(1-x) \) for \( 0 \leq x \leq 1 \) (and = 0 elsewhere); show that it is a valid density function. Find the cdf \( F(x) \), and sketch its graph. Calculate \( P(\frac{1}{4} \leq X \leq \frac{3}{4}) \).

Exercise: Sketch the graph of \( f(x) = \frac{e^x}{(e^x + 1)^2} \) for \(-\infty < x < +\infty\), and show that it is a valid density function. Find the cdf \( F(x) \), and sketch its graph. Find the quartiles. Calculate \( P(0 \leq X \leq 1) \).

If \( X \) is a *continuous* numerical random variable with **probability density function** (pdf) \( f(x) \), then the **population mean** is given by the “first moment”

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

and the **population variance** is given by the “second moment” about the mean

\[
\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) \, dx,
\]

or equivalently,

\[
\sigma^2 = E[X^2] - \mu^2 = \int_{-\infty}^{+\infty} x^2 f(x) \, dx - \mu^2.
\]

(Compare these *continuous* formulas with those for *discrete* \( X \).

Thus, for the exponential density, \( \mu = \int_0^\infty x a e^{-ax} \, dx = \frac{1}{a} \), via **integration by parts**.

The calculation of \( \sigma^2 \) is left as an exercise.

Exercise: Sketch the graph of \( f(x) = \frac{2}{\pi} \frac{1}{\sqrt{1-x^2}} \) for \( 0 \leq x < 1 \) (and 0 elsewhere); show that it is a valid density function. Find the cdf \( F(x) \), and sketch its graph. Calculate \( P(X \leq \frac{1}{2}) \), and find the mean.

Exercise: What are the mean and variance of the power density?

Exercise: What is the mean of the Cauchy density?

**Faites attention!**

**Ce n’est pas aussi facile qu’il apparaît...**

Augustin-Louis Cauchy

1789-1857
Example:

**Crawling Ants and Jumping Fleas**

Consider two insects on a (six-inch) ruler: a flea, who makes only discrete integer jumps \((X)\), and an ant, who crawls along continuously and can stop anywhere \((Y)\).

1. Let the discrete random variable \(X = \text{“length jumped (0, 1, 2, 3, 4, 5, or 6 inches) by the flea”} \). Suppose that the flea is tired, so is less likely to make a large jump than a small (or no) jump, according to the following probability mass function (pmf) \(p(x) = P(X = x)\), and corresponding probability histogram.

- The total probability is \(P(0 \leq X \leq 6) = 1\), as it should be.
- \(P(3 \leq X \leq 6) = 4/28 + 3/28 + 2/28 + 1/28 = 10/28\)
- \(P(0 \leq X < 3) = 7/28 + 6/28 + 5/28 = 18/28\), or
  \[= 1 - P(3 \leq X \leq 6) = 1 - 10/28 = 18/28\]
- \(P(0 \leq X \leq 3) = 18/28 + 4/28 = 22/28\), because \(P(X = 3) = 4/28\)

- **Exercise**: Confirm that the flea jumps a mean length of \(\mu = 2\) inches.
- **Exercise**: Sketch a graph of the cumulative distribution function \(F(x) = P(X \leq x)\), similar to that of §2.2 in these notes.
2. Let the continuous random variable \( Y = \text{“length crawled (any value in the interval [0, 6] inches) by the ant”} \). Suppose that the ant is tired, so is less likely to crawl a long distance than a short (or no) distance, according to the following probability density function (pdf) \( f(y) \), and its corresponding graph, the probability density curve. (Assume that \( f = 0 \) outside [0, 6].)

\[
f(y) = \frac{6 - y}{18}, \quad 0 \leq y \leq 6
\]

- The total probability is \( P(0 \leq Y \leq 6) = \frac{1}{2} \cdot 6 \cdot \frac{1}{3} = 1 \), as it should be.
- \( P(3 \leq Y \leq 6) = \frac{1}{2} \cdot 3 \cdot \frac{1}{6} = \frac{1}{4} \) (Could also use calculus.)

Equal!  
\[
\begin{align*}
P(0 \leq Y < 3) &= 1 - P(3 \leq Y \leq 6) = 1 - \frac{1}{4} = \frac{3}{4} \\
P(0 \leq Y \leq 3) &= \frac{3}{4} \text{ also, because } P(Y = 3) = 0 \quad \text{Why?}
\end{align*}
\]

- **Exercise:** Confirm that the ant crawls a mean length of \( \mu = 2 \) inches.
- **Exercise:** Find the cumulative distribution function \( F(y) \), and sketch its graph.
An extremely important bell-shaped continuous population distribution...

Normal Distribution (a.k.a. Gaussian Distribution): \( X \sim N(\mu, \sigma) \)

\[
f(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2}, \quad -\infty < x < +\infty
\]

\(
\pi = 3.14159...
\)

\(
e = 2.71828...
\)

Johann Carl Friedrich Gauss
(1777 - 1855)

\[
\text{Total Area } = \int_{-\infty}^{\infty} f(x) \, dx = 1
\]

Examples:

\[
\mu = 98.6
\]
\( X = \text{Body Temp (°F)} \)

\[
\mu = 100
\]
\( X = \text{IQ score (discrete!)} \)
Example: Two exams are given in a statistics course, both resulting in class scores that are normally distributed. The first exam distribution has a mean of 80.7 and a standard deviation of 3.5 points. The second exam distribution has a mean of 82.8 and a standard deviation of 4.5 points. Carla receives a score of 87 on the first exam, and a score of 90 on the second exam. Which of her two exam scores represents the better effort, relative to the rest of the class?

The \textbf{Z-score Transformation} tells how many standard deviations \( \sigma \) the \textbf{X-score} lies from the mean \( \mu \).

\[
X \sim N(\mu, \sigma) \iff Z = \frac{X - \mu}{\sigma} \sim N(0, 1)
\]

\textbf{Standard Normal Distribution}

\[
x \text{-score} = 87 \iff z \text{-score} = \frac{87 - 80.7}{3.5} = 1.8
\]

higher relative score

\[
x \text{-score} = 90 \iff z \text{-score} = \frac{90 - 82.8}{4.5} = 1.6
\]
Example: \(X = \text{“Age (years) of UW-Madison third-year undergraduate population”}\)

Assume: \(X \sim N(20, 1.25)\), i.e., \(X\) is \textbf{normally distributed} with mean \(\mu = 20\) yrs, and s.d. \(\sigma = 1.25\) yrs.

How do we check this? And what do we do if it’s not true, or we can’t tell? Later...

Suppose that an individual from this population is randomly selected. Then...

- \(P(X < 20) = 0.5\) (via symmetry)
- \(P(X < 19) = P\left( Z < \frac{19-20}{1.25} \right) = P(Z < -0.8) = 0.2119\) (via table or software)

Therefore...

- \(P(19 \leq X < 20) = P(X < 20) - P(X < 19) = 0.5000 - 0.2119 = 0.2881\)

Likewise,

- \(P(19 \leq X < 19.5) = 0.3446 - 0.2119 = 0.1327\)
- \(P(19 \leq X < 19.05) = 0.2236 - 0.2119 = 0.0118\)
- \(P(19 \leq X < 19.005) = 0.2130 - 0.2119 = 0.0012\)
- \(P(19 \leq X < 19.0005) = 0.2120 - 0.2119 = 0.0001\)
- \(P(X = 19.00000...) = 0,\) since \(X\) is \textit{continuous}!
Two Related Questions…

1. Given $X \sim N(\mu, \sigma)$. What is the probability that a randomly selected individual from the population falls within one standard deviation (i.e., $\pm \sigma$) of the mean $\mu$? Within two standard deviations ($\pm 2\sigma$)? Within three ($\pm 3\sigma$)?

Solution: We solve this by transforming to the tabulated standard normal distribution $Z \sim N(0, 1)$, via the formula $Z = \frac{X - \mu}{\sigma}$, i.e., $X = \mu + Z\sigma$.

\[
P(\mu - 1\sigma \leq X \leq \mu + 1\sigma) = P(-1 \leq Z \leq +1) = P(Z \leq +1) - P(Z \leq -1) = 0.8413 - 0.1587 = 0.6827
\]

\[
P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = P(-2 \leq Z \leq +2) = P(Z \leq +2) - P(Z \leq -2) = 0.9772 - 0.0228 = 0.9545
\]

Likewise, $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = P(-3 \leq Z \leq +3) = 0.9973$.

These so-called empirical guidelines can be used as an informal check to see if sample-generated data derive from a population that is normally distributed. For if so, then 68%, or approximately 2/3, of the data should lie within one standard deviation $s$ of the mean $\mu$; approximately 95% should lie within two standard deviations $2s$ of the mean $\mu$, etc. Other quantiles can be checked similarly. Superior methods also exist…

See my homepage to view a “ball drop” computer simulation of the normal distribution: (requires Java)

[http://www.stat.wisc.edu/~ifischer]
2. Given $X \sim N(\mu, \sigma)$. What symmetric interval about the mean $\mu$ contains 90% of the population distribution? 95%? 99%? General formulation?

Solution: Again, we can answer this question for the standard normal distribution $Z \sim N(0, 1)$, and transform back to $X \sim N(\mu, \sigma)$, via the formula $Z = \frac{X - \mu}{\sigma}$, i.e., $X = \mu + Z\sigma$.

The value $z_{0.05} = 1.645$ satisfies

$$P(-z_{0.05} \leq Z \leq z_{0.05}) = 0.90,$$

or equivalently,

$$P(Z \leq -z_{0.05}) = P(Z \geq z_{0.05}) = 0.05.$$

Hence, the required interval is $\mu - 1.645\sigma \leq X \leq \mu + 1.645\sigma$.

The value $z_{0.025} = 1.960$ satisfies

$$P(-z_{0.025} \leq Z \leq z_{0.025}) = 0.95,$$

or equivalently,

$$P(Z \leq -z_{0.025}) = P(Z \geq z_{0.025}) = 0.025.$$

Hence, the required interval is $\mu - 1.960\sigma \leq X \leq \mu + 1.960\sigma$.

The value $z_{0.005} = 2.575$ satisfies

$$P(-z_{0.005} \leq Z \leq z_{0.005}) = 0.99,$$

or equivalently,

$$P(Z \leq -z_{0.005}) = P(Z \geq z_{0.005}) = 0.005.$$

Hence, the required interval is $\mu - 2.575\sigma \leq X \leq \mu + 2.575\sigma$.

Def: The critical value $z_{0.025}$ satisfies

$$P(-z_{0.025} \leq Z \leq z_{0.025}) = 1 - \alpha,$$

or equivalently, the “tail probabilities”

$$P(Z \leq -z_{0.025}) = P(Z \geq z_{0.025}) = \alpha/2.$$

Hence, the required interval satisfies

$$P(\mu - z_{0.025}\sigma \leq X \leq \mu + z_{0.025}\sigma) = 1 - \alpha.$$
Normal Approximation to the Binomial Distribution

(continuous)  (discrete)

Example: Suppose that it is estimated that 20% (i.e., \( \pi = 0.2 \)) of a certain population has diabetes. Out of \( n = 100 \) randomly selected individuals, what is the probability that...

(a) exactly \( X = 10 \) are diabetics? \( X = 15? \ X = 20? \ X = 25? \ X = 30? \)

Assuming that the occurrence of diabetes is independent among the individuals in the population, we have \( X \sim \text{Bin}(100, 0.2) \). Thus, the values of \( P(X = x) \) are calculated in the following probability table and histogram.

<table>
<thead>
<tr>
<th>( x )</th>
<th>( P(X = x) = \binom{100}{x} (0.2)^x (0.8)^{100-x} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>( \binom{100}{10} (0.2)^{10} (0.8)^{90} = 0.00336 )</td>
</tr>
<tr>
<td>15</td>
<td>( \binom{100}{15} (0.2)^{15} (0.8)^{85} = 0.04806 )</td>
</tr>
<tr>
<td>20</td>
<td>( \binom{100}{20} (0.2)^{20} (0.8)^{80} = 0.09930 )</td>
</tr>
<tr>
<td>25</td>
<td>( \binom{100}{25} (0.2)^{25} (0.8)^{75} = 0.04388 )</td>
</tr>
<tr>
<td>30</td>
<td>( \binom{100}{30} (0.2)^{30} (0.8)^{70} = 0.00519 )</td>
</tr>
</tbody>
</table>

(b) \( X \leq 10 \) are diabetics? \( X \leq 15? \ X \leq 20? \ X \leq 25? \ X \leq 30? \)

Method 1: Directly sum the exact binomial probabilities to obtain \( P(X \leq x) \).

For instance, the cumulative probability \( P(X \leq 10) = \)

\[
\binom{100}{0} (0.2)^0 (0.8)^{100} + \binom{100}{1} (0.2)^1 (0.8)^{99} + \binom{100}{2} (0.2)^2 (0.8)^{98} + \binom{100}{3} (0.2)^3 (0.8)^{97} + \\
\binom{100}{4} (0.2)^4 (0.8)^{96} + \binom{100}{5} (0.2)^5 (0.8)^{95} + \binom{100}{6} (0.2)^6 (0.8)^{94} + \binom{100}{7} (0.2)^7 (0.8)^{93} + \\
\binom{100}{8} (0.2)^8 (0.8)^{92} + \binom{100}{9} (0.2)^9 (0.8)^{91} + \binom{100}{10} (0.2)^{10} (0.8)^{90} = 0.00570
\]
Method 2: Despite the skew, $X \sim N(\mu, \sigma)$, approximately (a consequence of the **Central Limit Theorem**, §5.2), with mean $\mu = n\pi$, and standard deviation $\sigma = \sqrt{n\pi(1 - \pi)}$. Hence,

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

becomes

$$Z = \frac{X - n\pi}{\sqrt{n\pi(1 - \pi)}} \sim N(0, 1).$$

In this example, 

$\mu = n\pi = (100)(0.2) = 20$, and 

$\sigma = \sqrt{n\pi(1 - \pi)} = \sqrt{100(0.2)(0.8)} = 4$.

So, approximately, $X \sim N(20, 4)$; thus 

$$Z = \frac{X - 20}{4} \sim N(0, 1).$$

For instance, $P(X \leq 10) \approx P\left(Z \leq \frac{10 - 20}{4}\right) = P(Z \leq -2.5) = 0.00621$.

The following table compares the two methods for finding $P(X \leq x)$.

<table>
<thead>
<tr>
<th></th>
<th>Binomial (exact)</th>
<th>Normal (approximation)</th>
<th>Normal (with correction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.00570</td>
<td>0.00621</td>
<td>0.00877</td>
</tr>
<tr>
<td>15</td>
<td>0.12851</td>
<td>0.10565</td>
<td>0.13029</td>
</tr>
<tr>
<td>20</td>
<td>0.55946</td>
<td>0.50000</td>
<td>0.54974</td>
</tr>
<tr>
<td>25</td>
<td>0.91252</td>
<td>0.89435</td>
<td>0.91543</td>
</tr>
<tr>
<td>30</td>
<td>0.99394</td>
<td>0.99379</td>
<td>0.99567</td>
</tr>
</tbody>
</table>

**Comment:** The **normal approximation to the binomial** generally works well, provided $n\pi \geq 15$ and $n(1 - \pi) \geq 15$. A modification exists, which adjusts for the difference between the discrete and continuous distributions:

$$Z = \frac{X - n\pi \pm 0.5}{\sqrt{n\pi(1 - \pi)}} \sim N(0, 1)$$

where the **continuity correction** factor is equal to $+0.5$ for $P(X \leq x)$, and $-0.5$ for $P(X \geq x)$. In this example, the “corrected” formula becomes 

$$Z = \frac{X - 20 + 0.5}{4} \sim N(0, 1).$$
Exercise: Recall the preceding section, where a spontaneous medical condition affects 1% (i.e., \( \pi = 0.01 \)) of the population, and \( X = \) “number of affected individuals in a random sample of \( n = 300 \).” Previously, we calculated the probability \( P(X = x) \) for \( x = 0, 1, \ldots, 300 \). We now ask for the more meaningful cumulative probability \( P(X \leq x) \), for \( x = 0, 1, 2, 3, 4, \ldots \). Rather than summing the exact binomial (or the approximate Poisson) probabilities as in Method 1 above, adopt the technique in Method 2, both with continuity correction and without. Compare these values with the exact binomial sums.

A Word about “Probability Zero” Events
(Much Ado About Nothing?)

Exactly what does it mean to say that an event \( E \) has zero probability of occurrence, i.e. \( P(E) = 0 \)? A common, informal interpretation of this statement is that the event “cannot happen” and, in many cases, this is indeed true. For example, if \( X = \) “Sum of two dice,” then “\( X = -4 \),” “\( X = 5.7 \),” and “\( X = 13 \)” all have probability zero because they are impossible outcomes of this experiment, i.e., they are not in the sample space \( \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\} \).

However, in a formal mathematical sense, this interpretation is too restrictive. For example, consider the following scenario: Suppose that \( k \) people participate in a lottery; each individual holds one ticket with a unique integer from the sample space \( \{1, 2, 3, \ldots, k\} \). The winner is determined by a computer that randomly selects one of these \( k \) integers with equal likelihood. Hence, the probability that a randomly selected individual wins is equal to \( 1/k \). The larger the number \( k \) of participants, the smaller the probability \( 1/k \) that any particular person will win. Now, for the sake of argument, suppose that there is an infinite number of participants; a computer randomly selects one integer from the sample space \( \{1, 2, 3, \ldots\} \). The probability that a randomly selected individual wins is therefore less than \( 1/k \) for any \( k \), i.e., arbitrarily small, hence \( = 0 \).* But by design, someone must win the lottery, so “probability zero” does not necessarily translate into “the event cannot happen.” So what does it mean?

Recall that the formal, classical definition of the probability \( P(E) \) of any event \( E \) is the mathematical “limiting value” of the ratio \( \frac{\#(E \text{ occurs})}{\# \text{ trials}} \), as \( \# \text{ trials} \to \infty \). That is, the fraction of “the number of times that the event occurs” to “the total number of experimental trials,” as the experiment is repeated indefinitely. If, in principle, this ratio becomes arbitrarily small after sufficiently many trials, then such an ever-increasingly rare event \( E \) is formally identified with having “probability zero” (such as, perhaps, the random toss of a coin under ordinary conditions resulting in it landing on edge, rather than on heads or tails).

* Similarly, any event consisting of a finite subset of an infinite sample space of possible outcomes (such as the event of randomly selecting a single particular value from a continuous interval), has a mathematical probability of zero.
Ismor Fischer, 5/26/2016

Classical Continuous Probability Densities
(The \( t \) and \( F \) distributions will be handled separately.)

**Uniform**

\[ f(x) = \frac{1}{b-a}, \quad a \leq x \leq b \]

Consequently, \( F(x) = \frac{x-a}{b-a} \).

**Normal**

For \( \sigma > 0 \),

\[ f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < +\infty. \]

**Log-Normal**

For \( \beta > 0 \),

\[ f(x) = \frac{1}{\sqrt{2\pi}\beta} x^{-\frac{1}{2}} e^{-\frac{1}{2}\left(\frac{\ln x-\alpha}{\beta}\right)^2}, \quad x \geq 0. \]

**Gamma**

For \( \alpha > 0 \), \( \beta > 0 \),

\[ f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, \quad x \geq 0. \]

**Chi-Squared**: For \( \nu = 1, 2, \ldots \)

\[ f(x) = \frac{1}{2^{\nu/2} \Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2}, \quad x \geq 0. \]

**Beta**

For \( \alpha > 0 \), \( \beta > 0 \),

\[ f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad 0 \leq x \leq 1. \]

**Weibull**

**Exponential**

For \( \alpha > 0 \), \( \beta > 0 \),

\[ f(x) = \frac{1}{\beta} e^{-x/\beta}, \quad x \geq 0. \]

Thus, \( F(x) = 1 - e^{-x/\beta} \).

**Notes on the Gamma and Beta Functions**

**Def**: \( \Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \, dx \)

**Thm**: \( \Gamma(\alpha) = (\alpha-1) \Gamma(\alpha-1) \); therefore, \( = (\alpha-1)!, \) if \( \alpha = 1, 2, 3, \ldots \)

**Thm**: \( \Gamma(1/2) = \sqrt{\pi} \)

**Def**: \( B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} \, dx \)

**Thm**: \( B(\alpha, \beta) = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha+\beta)} \)