Section 5.1 Continuous Random Variables: Introduction

Not all random variables are discrete. For example:

1. Waiting times for anything (train, arrival of customer, production of mRNA molecule from gene, etc).

2. Distance a ball is thrown.

3. Size of an antenna on a bug.

The general idea is that now the sample space is *uncountable*. Probability mass functions and summation no longer works.

DEFINITION: We say that X is a *continuous* random variable if the sample space is uncountable and there exists a nonnegative function f defined for all $x \in (-\infty, \infty)$ having the property that for any set $B \subset \mathbb{R}$,

$$\boxed{P\{X \in B\} = \int_B f(x)dx}$$

The function f is called the *probability density function* of the random variable X, and is (sort of) the analogue of the probability mass function in the discrete case.

So probabilities are now found by integration, rather than summation.

REMARK: We must have that

$$1 = P\{-\infty < X < \infty\} = \int_{-\infty}^{\infty} f(x)dx$$

Also, taking $B = [a, b]$ for $a < b$ we have

$$\boxed{P\{a \le X \le b\} = \int_a^b f(x)dx}$$

Note that taking $a = b$ yields the moderately counter-intuitive

$$P\{X = a\} = \int_a^a f(x)dx = 0$$

Recall: For discrete random variables the probability mass function can be reconstructed from the distribution function and vice versa and changes in the distribution function corresponded with finding where the "mass" of the probability was.

We now have that

$$\boxed{F(t) = P\{X \le t\} = \int_{-\infty}^{t} f(x)dx}$$

and so

$$\boxed{F'(t) = f(t)}$$

agreeing with the previous interpretation. Also,

$$\boxed{P(X \in (a,\ b)) = P(X \in [a,\ b]) = etc. \ = \int_a^b f(t)dt = F(b) - F(a)}$$

The density function does not represent a probability. However, its integral gives probability of being in certain regions of $\mathbb{R}$. Also, $f(a)$ gives a measure of likelihood of being around a. That is,

$$P\{a - \epsilon/2 < X < a + \epsilon/2\} = \int_{a-\epsilon/2}^{a+\epsilon/2} f(t)dt \approx f(a)\epsilon$$

when ϵ is small and f is continuous at a. Thus, the probability that X will be in an interval around a of size ϵ is approximately $\epsilon f(a)$. So, if $f(a) < f(b)$, then

$$P(a - \epsilon/2 < X < a + \epsilon/2) \approx \epsilon f(a) < \epsilon f(b) \approx P(b - \epsilon/2 < X < b + \epsilon/2)$$

In other words, $f(a)$ is a measure of how likely it is that the random variable takes a value near a.

EXAMPLES:

1. The amount of time you must wait, in minutes, for the appearance of an mRNA molecule is a continuous random variable with density

$$f(t) = \begin{cases} \lambda e^{-3t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

What is the probability that

1. You will have to wait between 1 and 2 minutes?
2. You will have to wait longer than 1/2 minutes?

Solution: First, we haven't given λ. We need

$$1 = \int_{-\infty}^{\infty} f(t)dt = \int_0^{\infty} \lambda e^{-3t}dt = -\frac{\lambda}{3}e^{-3t}|_{t=0}^{\infty} = \frac{\lambda}{3}$$

Therefore, $\lambda = 3$, and $f(t) = 3e^{-3t}$ for $t \geq 0$. Thus,

$$P\{1 < X < 2\} = \int_1^2 3e^{-3t}dt = e^{-3\cdot 1} - e^{-3\cdot 2} \approx 0.0473$$

$$P\{X > 1/2\} = \int_{1/2}^{\infty} 3e^{-3t}dt = e^{-3/2} \approx 0.2231$$

2. If X is a continuous random variable with distribution function F_X and density f_X, find the density function of $Y = kX$.

2. If X is a continuous random variable with distribution function F_X and density f_X, find the density function of $Y = kX$.

Solution: We'll do two ways (like in the book). We have

$$\begin{aligned} F_Y(t) &= P\{Y \le t\} \\ &= P\{kX \le t\} \\ &= P\{X \le t/k\} \\ &= F_X(t/k) \end{aligned}$$

Differentiating with respect to t yields

$$f_Y(t) = \frac{d}{dt} F_X(t/k) = \frac{1}{k} f_X(t/k)$$

Another derivation is the following. We have

$$\begin{aligned} \epsilon f_Y(a) &\approx P\{a - \epsilon/2 \le Y \le a + \epsilon/2\} \\ &= P\{a - \epsilon/2 \le kX \le a + \epsilon/2\} \\ &= P\{a/k - \epsilon/(2k) \le X \le a/k + \epsilon/(2k)\} \\ &\approx \frac{\epsilon}{k} f_X(a/k) \end{aligned}$$

Dividing by ϵ yields the same result as before.

Returning to our previous example with $f_X(t) = 3e^{-3t}$. If $Y = 3X$, then

$$f_Y(t) = \frac{1}{3} f_X(t/3) = e^{-t}, \ t \ge 0.$$

What if $Y = kX + b$? We have

$$\begin{aligned} F_Y(t) &= P\{Y \le t\} \\ &= P\{kX + b \le t\} \\ &= P\{X \le (t - b)/k\} \\ &= F_X\left(\frac{t - b}{k}\right) \end{aligned}$$

Differentiating with respect to t yields

$$f_Y(t) = \frac{d}{dt} F_X\left(\frac{t - b}{k}\right) = \frac{1}{k} f_X\left(\frac{t - b}{k}\right)$$

Section 5.2 Expectation and Variance of Continuous Random Variables

DEFINITION: If X is a random variable with density function f, then the expected value is

$$\boxed{E[X]=\int_{-\infty}^{\infty} xf(x)dx}$$

This is analogous to the discrete case:

1. Discretize X into small ranges $(x_{i-1},\ x_i]$ where $x_i - x_{i-1} = h$ is small.
2. Now think of X as discrete with the values x_i.
3. Then,

$$E[X]\approx\sum_{x_i} x_i p(x_i)\approx\sum_{x_i} x_i P(x_{i-1}<X\leq x_i)\approx\sum_{x_i} x_i f(x_i)h\approx\int_{-\infty}^{\infty} xf(x)dx$$

EXAMPLES:

1. Suppose that X has the density function

$$f(x)=\frac{1}{2}x,\ \ 0\leq x\leq 2$$

Find $E[X]$.

Solution: We have

$$E[X]=\int_0^2 x\cdot\frac{1}{2}xdx=\frac{1}{6}x^3\bigg|_0^2=\frac{8}{6}=\frac{4}{3}$$

2. Suppose that $f_X(x)=1$ for $x\in(0,1)$. Find $E[e^X]$.

2. Suppose that $f_X(x) = 1$ for $x \in (0,1)$. Find $E[e^X]$.

Solution: Let $Y = e^X$. We need to find the density of Y, then we can use the definition of Y. Because the range of X is $(0,1)$, the range of Y is $(1,e)$. For $1 \le x \le e$ we have

$$F_Y(x) = P\{Y \le x\} = P\{e^X \le x\} = P\{X \le \ln x\} = F_X(\ln x) = \int_0^{\ln x} f_X(t)dt = \ln x$$

Differentiating both sides, we get $f_Y(x) = 1/x$ for $1 \le x \le e$ and zero otherwise. Thus,

$$E[e^X] = E[Y] = \int_{-\infty}^{\infty} x f_Y(x)dx = \int_1^e dx = e - 1$$

As in the discrete case, there is a theorem making these computations much easier.

THEOREM: Let X be a continuous random variable with density function f. Then for any $g : \mathbb{R} \to \mathbb{R}$ we have

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

Note how easy this makes the previous example:

$$E[e^X] = \int_0^1 e^x dx = e - 1$$

To prove the theorem in the special case that $g(x) \ge 0$ we need the following:

LEMMA: For a nonnegative random variable Y,

$$E[Y] = \int_0^{\infty} P\{Y > y\}dy$$

Proof: We have

$$\int_0^{\infty} P\{Y > y\}dy = \int_0^{\infty}\int_y^{\infty} f_Y(x)dxdy = \int_0^{\infty}\int_0^{x} f_Y(x)dydx$$

$$= \int_0^{\infty} f_Y(x)\int_0^{x} dydx = \int_0^{\infty} x f_Y(x)dx = E[Y] \blacksquare$$

Proof of Theorem: For any function $g : \mathbb{R} \to \mathbb{R}_{\ge 0}$ (general case is similar) we have

$$E[g(X)] = \int_0^{\infty} P\{g(X) > y\}dy = \int_0^{\infty}\left[\int_{x:g(x)>y} f(x)dx\right]dy$$

$$= \iint_{\{(x,y)\}:g(x)>y\ge 0} f(x)dxdy = \int_{x:g(x)>0} f(x)\int_0^{g(x)} dydx$$

$$= \int_{x:g(x)>0} f(x)g(x)dx = \int_{-\infty}^{\infty} f(x)g(x)dx \blacksquare$$

Of course, it immediately follows from the above theorem that for any constants a and b we have

$$\boxed{E[aX+b]=aE[X]+b}$$

In fact, putting $g(x)=ax+b$ in the previous theorem, we get

$$E[aX+b]=\int_{-\infty}^{\infty}(ax+b)f_X(x)dx=a\int_{-\infty}^{\infty}xf_X(x)dx+b\int_{-\infty}^{\infty}f_X(x)dx=aE[X]+b$$

Thus, expected values inherit their linearity from the linearity of the integral.

DEFINITION: If X is random variable with mean μ, then the variance and standard deviation are given by

$$\boxed{\begin{gathered}\text{Var}(X)=E[(X-\mu)^2]=\int_{-\infty}^{\infty}(x-\mu)^2f(x)dx\\ \sigma_X=\sqrt{\text{Var}(x)}\end{gathered}}$$

We also still have

$$\boxed{\begin{gathered}\text{Var}(X)=E[X^2]-(E[X])^2\\ \text{Var}(aX+b)=a^2\text{Var}(X)\\ \sigma_{aX+b}=|a|\sigma_X\end{gathered}}$$

The proofs are exactly the same as in the discrete case.

EXAMPLE: Consider again X with the density function

$$f(x)=\frac{1}{2}x,\ 0\le x\le 2$$

Find Var$[X]$.

Solution: Recall that we have $E[X]=4/3$. We now find the second moment

$$E[X^2]=\int_{-\infty}^{\infty}x^2f(x)dx=\int_0^2 x^2\frac{1}{2}xdx=\frac{1}{8}x^4\Big|_0^2=\frac{16}{8}=2$$

Therefore,

$$\text{Var}(X)=E[X^2]-E[X]^2=2-(4/3)^2=18/9-16/9=1/3$$

Section 5.3 The Uniform Random Variable

We first consider the interval $(0,1)$ and say a random variable is *uniformly* distributed over $(0,1)$ if its density function is

$$\boxed{f(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \text{else} \end{cases}}$$

Note that

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^1 dx = 1$$

Also, as $f(x) = 0$ outside of $(0,1)$, X only takes values in $(0,1)$.

As f is a constant, X is equally likely to be near each number. That is, each small region of size ϵ is equally likely to contain X (and has a probability of ϵ of doing so).

Note that for any $0 \le a \le b \le 1$ we have

$$\boxed{P\{a \le X \le b\} = \int_a^b f(x)dx = b - a}$$

General Case: Now consider an interval (a,b). We say that X is *uniformly* distributed on (a,b) if

$$\boxed{f(t) = \begin{cases} \dfrac{1}{b-a}, & a < t < b \\ 0, & \text{else} \end{cases}} \quad \text{and} \quad \boxed{F(t) = \begin{cases} 0, & t < a \\ \dfrac{t-a}{b-a}, & a \le x < b \\ 1, & t \ge b \end{cases}}$$

We can compute the expected value and variance straightaway:

We can compute the expected value and variance straightaway:

$$E[X] = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \cdot \frac{1}{2}(b^2 - a^2) = \boxed{\frac{b+a}{2}}$$

Similarly,

$$E[X^2] = \int_a^b x^2 \cdot \frac{1}{b-a} dx = \frac{b^3 - a^3}{3(b-a)} = \frac{b^2 + ab + a^2}{3}$$

Therefore,

$$\text{Var}(X) = E[X^2] - E[X]^2 = \frac{b^2 + ab + a^2}{3} - \left(\frac{b+a}{2}\right)^2 = \boxed{\frac{(b-a)^2}{12}}$$

EXAMPLE: Consider a uniform random variable with density

$$f(x) = \begin{cases} \frac{1}{8}, & x \in (-1, 7) \\ 0, & \text{else} \end{cases}$$

Find $P\{X < 2\}$.

EXAMPLE: Consider a uniform random variable with density

$$f(x) = \begin{cases} \frac{1}{8}, & x \in (-1, 7) \\ 0, & \text{else} \end{cases}$$

Find $P\{X < 2\}$.

Solution: The range of X is $(-1, 7)$. Thus, we have

$$P\{X < 2\} = \int_{-\infty}^{2} f(x)dx = \int_{-1}^{2} \frac{1}{8}dx = \frac{3}{8}$$

Section 5.4 Normal Random Variables

We say that X is a *normal random variable* or simply that X is *normal* with parameters μ and σ^2 if the density is given by

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

This density is a bell shaped curve and is symmetric around μ:

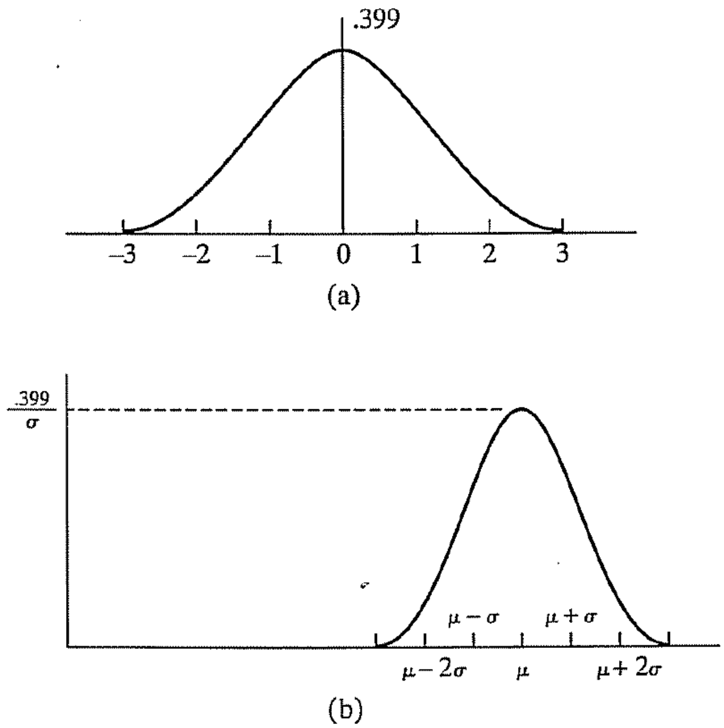

Figure 5.5 Normal density function: (a) $\mu = 0$, $\sigma = 1$; (b) arbitrary μ, σ^2.

It should NOT be apparent that this is a density. We need that it integrates to one. Making the substitution $y = (x - \mu)/\sigma$, we have

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy$$

We will show that the integral is equal to $\sqrt{2\pi}$. To do so, we actually compute the square of the integral. Define

$$I = \int_{-\infty}^{\infty} e^{-y^2/2} dy$$

Then,

$$I^2 = \int_{-\infty}^{\infty} e^{-y^2/2} dy \int_{-\infty}^{\infty} e^{-x^2/2} dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2} dx dy$$

Now we switch to polar coordinates. That is,

$$x = r\cos\theta, \quad y = r\sin\theta, \quad dxdy = rd\theta dr$$

Thus,

$$I^2 = \int_0^{\infty}\int_0^{2\pi} e^{-r^2/2} r d\theta dr = \int_0^{\infty} e^{-r^2/2} r \int_0^{2\pi} d\theta dr = \int_0^{\infty} 2\pi e^{-r^2/2} r dr = -2\pi e^{-r^2/2}\Big|_{r=0}^{\infty} = 2\pi$$

Scaling a normal random variable gives you another normal random variable.

THEOREM: Let X be Normal(μ, σ^2). Then $Y = aX + b$ is a normal random variable with parameters $a\mu + b$ and $a^2\sigma^2$

Proof: We let $a > 0$ (the proof for $a < 0$ is the same). We first consider the distribution function of Y. We have

$$F_Y(t) = P\{Y \le t\} = P\{aX + b \le t\} = P\{X \le (t-b)/a\} = F_X\left(\frac{t-b}{a}\right)$$

Differentiating both sides, we get

$$f_Y(t) = \frac{1}{a} f_X\left(\frac{t-b}{a}\right) = \frac{1}{\sigma a\sqrt{2\pi}} \exp\left\{ -\frac{\left(\frac{t-b}{a} - \mu\right)^2}{2\sigma^2} \right\}$$

$$= \frac{1}{\sigma a\sqrt{2\pi}} \exp\left\{ -\frac{(t-b-a\mu)^2}{2(a\sigma)^2} \right\}$$

which is the density of a normal random variable with parameters $a\mu + b$ and $a^2\sigma^2$. ■

From the above theorem it follows that if X is Normal(μ, σ^2), then

$$\boxed{Z = \frac{X - \mu}{\sigma}}$$

is normally distributed with parameters 0 and 1. Such a random variable is called a *standard normal* random variable. Note that its density is simply

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

By the above scaling, we see that we can compute the mean and variance for *all* normal random variables if we can compute it for the standard normal. In fact, we have

$$E[Z] = \int_{-\infty}^{\infty} x f_Z(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-x^2/2} dx = -\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \Big|_{-\infty}^{\infty} = 0$$

and

$$\text{Var}(Z) = E[Z^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx$$

Integration by parts with $u = x$ and $dv = xe^{-x^2/2}dx$ yields

$$\text{Var}(Z) = 1$$

Therefore, we see that if $X = \mu + \sigma Z$, then X is a normal random variable with parameters μ, σ^2 and

$$\boxed{E[X] = \mu, \quad \text{Var}(X) = \sigma^2 \text{Var}(Z) = \sigma^2}$$

NOTATION: Typically, we denote $\Phi(x) = P\{Z \le x\}$ for a standard normal Z. That is,

$$\boxed{\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy}$$

Note that by the symmetry of the density we have

$$\Phi(-x) = 1 - \Phi(x)$$

Therefore, it is customary to provide charts with the values of $\Phi(x)$ for all $x \in (0, 3.5)$ in increments of 0.01. (Page 201 in the book).

Note that if $X \sim$ Normal(μ, σ^2), then Φ can still be used:

$$\boxed{F_X(a) = P\{X \le a\} = P\{(X-\mu)/\sigma \le (a-\mu)/\sigma\} = \Phi\left(\frac{a-\mu}{\sigma}\right)}$$

EXAMPLE: Let $X \sim$ Normal$(3, 9)$, i.e. $\mu = 3$ and $\sigma^2 = 9$. Find $P\{2 < X < 5\}$.

EXAMPLE: Let $X \sim$ Normal$(3,9)$, i.e. $\mu=3$ and $\sigma^2=9$. Find $P\{2<X<5\}$.

Solution: We have, where Z is a standard normal,

$$P\{2<X<5\}=P\left\{\frac{2-3}{3}<Z<\frac{5-3}{3}\right\}=P\{-1/3<Z<2/3\}=\Phi(2/3)-\Phi(-1/3)$$

$$=\Phi(2/3)-(1-\Phi(1/3))\approx\Phi(0.66)-1+\Phi(0.33)=0.7454-1+0.6293=0.3747$$

We can use the normal random variable to approximate a binomial. This is a special case of the Central limit theorem.

THEOREM (The DeMoivre-Laplace limit theorem): Let S_n denote the number of successes that occur when n independent trials, each with a probability of p, are performed. Then for any $a<b$ we have

$$\boxed{P\left\{a\le\frac{S_n-np}{\sqrt{np(1-p)}}\le b\right\}\to\Phi(b)-\Phi(a)}$$

A rule of thumb is that this is a "good" approximation when $np(1-p)\ge 10$. Note that there is no assumption of p being small like in Poisson approximation.

EXAMPLE: Suppose that a biased coin is flipped 70 times and that probability of heads is 0.4. Approximate the probability that there are 30 heads. Compare with the actual answer. What can you say about the probability that there are between 20 and 40 heads?

EXAMPLE: Suppose that a biased coin is flipped 70 times and that probability of heads is 0.4. Approximate the probability that there are 30 heads. Compare with the actual answer. What can you say about the probability that there are between 20 and 40 heads?

Solution: Let X be the number of times the coin lands on heads. We first note that a normal random variable is continuous, whereas the binomial is discrete. Thus, (and this is called the *continuity correction*), we do the following

$$\begin{aligned}
P\{X = 30\} &= P\{29.5 < X < 30.5\} \\
&= P\left\{\frac{29.5 - 28}{\sqrt{70 \cdot 0.4 \cdot 0.6}} < \frac{X - 28}{\sqrt{70 \cdot 0.4 \cdot 0.6}} < \frac{30.5 - 28}{\sqrt{70 \cdot 0.4 \cdot 0.6}}\right\} \\
&\approx P\{0.366 < Z < 0.61\} \\
&\approx \Phi(0.61) - \Phi(0.37) \\
&\approx 0.7291 - 0.6443 \\
&= 0.0848
\end{aligned}$$

Whereas

$$P\{X = 30\} = \binom{70}{30} 0.4^{30} 0.6^{40} = 0.0853$$

We now evaluate the probability that there are between 20 and 40 heads:

$$\begin{aligned}
P\{20 < X < 40\} &= P\{19.5 < X < 40.5\} \\
&= P\left\{\frac{19.5 - 28}{\sqrt{70 \cdot 0.4 \cdot 0.6}} < \frac{X - 28}{\sqrt{70 \cdot 0.4 \cdot 0.6}} < \frac{40.5 - 28}{\sqrt{70 \cdot 0.4 \cdot 0.6}}\right\} \\
&\approx P\{-2.07 < Z < 3.05\} \\
&\approx \Phi(3.05) - \Phi(-2.07) \\
&= \Phi(3.05) - (1 - \Phi(2.07)) \\
&\approx 0.9989 - (1 - 0.9808) \\
&= 0.9797
\end{aligned}$$

Section 5.5 Exponential Random Variables

Exponential random variables arise as the distribution of the amount of time until some specific event occurs.

A continuous random variable with the density function

$$\boxed{f(x) = \lambda e^{-\lambda x} \quad \text{for} \quad x \geq 0}$$

and zero otherwise for some $\lambda > 0$ is called an *exponential* random variable with parameter $\lambda > 0$. The distribution function for $t \geq 0$ is

$$\boxed{F(t) = P\{X \leq t\} = \int_0^t \lambda e^{-\lambda x} dx = -e^{-\lambda x}\Big|_{x=0}^{t} = 1 - e^{-\lambda t}}$$

Computing the moments of an exponential is easy with integration by parts for $n > 0$:

$$E[X^n] = \int_0^\infty x^n \lambda e^{-\lambda x} dx = -x^n e^{-\lambda x}\Big|_{x=0}^{\infty} + \int_0^\infty n x^{n-1} e^{-\lambda x} dx = \frac{n}{\lambda}\int_0^\infty x^{n-1}\lambda e^{-\lambda x} dx = \frac{n}{\lambda} E[X^{n-1}]$$

Therefore, for $n = 1$ and $n = 2$ we see that

$$\boxed{E[X] = \frac{1}{\lambda}, \quad E[X^2] = \frac{2}{\lambda}E[X] = \frac{2}{\lambda^2}, \ldots, \ E[X^n] = \frac{n!}{\lambda^n}} \quad \text{and} \quad \boxed{\text{Var}(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}}$$

EXAMPLE: Suppose the length of a phone call in minutes is an exponential random variable with parameter $\lambda = 1/10$ (so the average call is 10 minutes). What is the probability that a random call will take

(a) more than 8 minutes?

(b) between 8 and 22 minutes?

Solution: Let X be the length of the call. We have

$$P\{X > 8\} = 1 - F(8) = 1 - \left(1 - e^{-(1/10)\cdot 8}\right) = e^{-0.8} = 0.4493$$

$$P\{8 < X < 22\} = F(22) - F(8) = e^{-0.8} - e^{-2.2} = 0.3385$$

Memoryless property: We say that a nonnegative random variable X is *memoryless* if

$$\boxed{P\{X > s + t \mid X > t\} = P\{X > s\} \quad \text{for all } s, t \geq 0}$$

Note that this property is similar to the geometric random variable one.

Let us show that an an exponential random variable satisfies the memoryless property. In fact, we have

$$P\{X > s + t \mid X > t\} = \frac{P\{X > s + t\}}{P\{X > t\}} = e^{-\lambda(s+t)}/e^{-\lambda t} = e^{-\lambda s} = P\{X > s\} \ \blacksquare$$

One can show that exponential is the only continuous distribution with this property.

EXAMPLE: Three people are in a post-office. Persons 1 and 2 are at the counter and will leave after an Exp(λ) amount of time. Once one leaves, person 3 will go to the counter and be served after an Exp(λ) amount of time. What is the probability that person 3 is the last to leave the post-office?

Solution: After one leaves, person 3 takes his spot. Next, by the memoryless property, the waiting time for the remaining person is again Exp(λ), just like person 3. By symmetry, the probability is then 1/2.

Hazard Rate Functions. Let X be a positive, continuous random variable that we think of as the lifetime of something. Let it have distribution function F and density f. Then the *hazard* or *failure rate* function $\lambda(t)$ of the random variable is defined via

$$\lambda(t) = \frac{f(t)}{\overline{F}(t)} \quad \text{where} \quad \overline{F} = 1 - F$$

What is this about? Note that

$$P\{X \in (t,\ t+dt) \mid X > t\} = \frac{P\{X \in (t, t+dt), X > t\}}{P\{X > t\}} = \frac{P\{X \in (t, t+dt)\}}{P\{X > t\}} \approx \frac{f(t)}{\overline{F}(t)} dt$$

Therefore, $\lambda(t)$ represents the conditional probability intensity that a t-unit-old item will fail.

Note that for the exponential distribution we have

$$\lambda(t) = \frac{f(t)}{\overline{F}(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$$

The parameter λ is usually referred to as the *rate* of the distribution.

It also turns out that the hazard function $\lambda(t)$ *uniquely* determines the distribution of a random variable. In fact, we have

$$\lambda(t) = \frac{\frac{d}{dt}F(t)}{1 - F(t)}$$

Integrating both sides, we get

$$\ln(1 - F(t)) = -\int_0^t \lambda(s)ds + k$$

Thus,

$$1 - F(t) = e^k \exp\left\{-\int_0^t \lambda(s)ds\right\}$$

Setting $t = 0$ (note that $F(0) = 0$) shows that $k = 0$. Therefore

$$F(t) = 1 - \exp\left\{-\int_0^t \lambda(s)ds\right\}$$

EXAMPLE: Suppose that the life distribution of an item has the hazard rate function

$$\lambda(t) = t^3, \quad t > 0$$

What is the probability that

(a) the item survives to age 2?

(b) the item's lifetime is between 0.4 and 1.4?

(c) the item survives to age 1?

(d) the item that survived to age 1 will survive to age 2?

EXAMPLE: Suppose that the life distribution of an item has the hazard rate function

$$\lambda(t) = t^3, \quad t > 0$$

What is the probability that

(a) the item survives to age 2?

(b) the item's lifetime is between 0.4 and 1.4?

(c) the item survives to age 1?

(d) the item that survived to age 1 will survive to age 2?

Solution: We are given $\lambda(t)$. We know that

$$F(t) = 1 - \exp\left\{-\int_0^t \lambda(s)ds\right\} = 1 - \exp\left\{-\int_0^t s^3 ds\right\} = 1 - \exp\left\{-t^4/4\right\}$$

Therefore, letting X be the lifetime of an item, we get

(a) $P\{X > 2\} = 1 - F(2) = \exp\{-2^4/4\} = e^{-4} \approx 0.01832$

(b) $P\{0.4 < X < 1.4\} = F(1.4) - F(0.4) = \exp\{-(0.4)^4/4\} - \exp\{-(1.4)^4/4\} \approx 0.6109$

(c) $P\{X > 1\} = 1 - F(1) = \exp\{-1/4\} \approx 0.7788$

(d) $P\{X > 2 \mid X > 1\} = \dfrac{P\{X > 2\}}{P\{X > 1\}} = \dfrac{1 - F(2)}{1 - F(1)} = \dfrac{e^{-4}}{e^{-1/4}} \approx 0.0235$

EXAMPLE: A smoker is said to have a rate of death "twice that" of a non-smoker at every age. What does this mean?

Solution: Let $\lambda_s(t)$ denote the hazard of a smoker and $\lambda_n(t)$ be the hazard of a non-smoker. Then

$$\lambda_s(t) = 2\lambda_n(t)$$

Let X_n and X_s denote the lifetime of a non-smoker and smoker, respectively. The probability that a non-smoker of age A will reach age $B > A$ is

$$P\{X_n > B \mid X_n > A\} = \frac{1 - F_n(B)}{1 - F_n(A)} = \frac{\exp\left\{-\int_0^B \lambda_n(t)dt\right\}}{\exp\left\{-\int_0^A \lambda_n(t)dt\right\}} = \exp\left\{-\int_A^B \lambda_n(t)dt\right\}$$

Whereas, the corresponding probability for a smoker is

$$P\{X_s > B \mid X_s > A\} = \exp\left\{-\int_A^B \lambda_s(t)dt\right\} = \exp\left\{-2\int_A^B \lambda_n(t)dt\right\} = \left[\exp\left\{-\int_A^B \lambda_n(t)dt\right\}\right]^2$$

$$= [P\{X_n > B \mid X_n > A\}]^2$$

Thus, if the probability that a 50 year old non-smoker reaches 60 is 0.7165, the probability for a smoker is $0.7165^2 = 0.5134$. If the probability a 50 year old non-smoker reaches 80 is 0.2, the probability for a smoker is $0.2^2 = 0.04$.

Section 5.6 Other Continuous Distributions

The Gamma distribution: A random variable X is said to have a gamma distribution with parameters $(\alpha,\ \lambda)$ for $\lambda > 0$ and $\alpha > 0$ if its density function is

$$f(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(x)}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

where $\Gamma(x)$ is called the *gamma function* and is defined by

$$\Gamma(x) = \int_0^\infty e^{-y} y^{x-1} dy$$

Note that $\Gamma(1) = 1$. For $x \neq 1$ we use integration by parts with $u = y^{x-1}$ and $dv = e^{-y}dy$ to show that

$$\Gamma(x) = -e^{-y}y^{x-1}|_{y=0}^{\infty} + (x-1)\int_0^\infty e^{-y}y^{x-2}dy = (x-1)\Gamma(x-1)$$

Therefore, for integer values of $n \geq 1$ we see that

$$\Gamma(n) = (n-1)\Gamma(n-1) = (n-1)(n-2)\Gamma(n-2) = \ldots = (n-1)(n-2)\ldots 3\cdot 2\Gamma(1) = (n-1)!$$

So, it is a generalization of the factorial.

REMARK: One can show that

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \quad \Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}, \quad \Gamma\left(\frac{5}{2}\right) = \frac{3}{4}\sqrt{\pi}, \quad \ldots, \quad \Gamma\left(\frac{1}{2}+n\right) = \frac{(2n)!}{4^n n!}\sqrt{\pi},\ n \in \mathbb{Z}_{\geq 0}$$

The gamma distribution comes up a lot as the sum of independent exponential random variables of parameter λ. That is, if $X_1, \ldots, X_n$ are independent Exp(λ), then $T_n = X_1 + \ldots + X_n$ is gamma(n, λ).

Recall that the time of the nth jump in a Poisson process, $N(t)$, is the sum of n exponential random variables of parameter λ. Thus, the time of the nth event follows a gamma$(n,\ \lambda)$ distribution. To prove all this we denote the time of the nth jump as T_n and see that

$$P\{T_n \leq t\} = P\{N(t) \geq n\} = \sum_{j=n}^{\infty} P\{N(t) = j\} = \sum_{j=n}^{\infty} e^{-\lambda t}\frac{(\lambda t)^j}{j!}$$

Differentiating both sides, we get

$$f_{T_n}(t) = \sum_{j=n}^{\infty} e^{-\lambda t}\frac{j(\lambda t)^{j-1}\lambda}{j!} - \lambda\sum_{j=n}^{\infty} e^{-\lambda t}\frac{(\lambda t)^j}{j!} = \sum_{j=n}^{\infty} \lambda e^{-\lambda t}\frac{(\lambda t)^{j-1}}{(j-1)!} - \lambda\sum_{j=n}^{\infty} e^{-\lambda t}\frac{(\lambda t)^j}{j!} = \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}$$

which is the density of a gamma(n, λ) random variable.

REMARK: Note that as expected gamma$(1, \lambda)$ is Exp(λ).

We have

$$E[X]=\frac{1}{\Gamma(x)}\int_0^\infty \lambda x e^{-\lambda x}(\lambda x)^{\alpha-1}dx=\frac{1}{\lambda\Gamma(x)}\int_0^\infty \lambda e^{-\lambda x}(\lambda x)^{\alpha}dx=\frac{\Gamma(\alpha+1)}{\lambda\Gamma(x)}=\frac{\alpha}{\lambda}$$

so

$$\boxed{E[X]=\frac{\alpha}{\lambda}}$$

Recall that the expected value of Exp(λ) is $1/\lambda$ and this ties in nicely with the sum of exponentials being a gamma. It can also be shown that

$$\boxed{\mathrm{Var}(X)=\frac{\alpha}{\lambda^2}}$$

DEFINITION: The gamma distribution with $\lambda=1/2$ and $\alpha=n/2$ where n is a positive integer is called the χ_n^2 *distribution* with n degrees of freedom.

Section 5.7 The Distribution of a Function of a Random Variable

We formalize what we've already done.

EXAMPLE: Let X be uniformly distributed over $(0,1)$. Let $Y=X^n$. For $0\le y\le 1$ we have

$$F_Y(y)=P\{Y\le y\}=P\{X^n\le y\}=P\{X\le y^{1/n}\}=F_X(y^{1/n})=y^{1/n}$$

Therefore, for $y\in(0,1)$ we have

$$f_Y(y)=\frac{1}{n}y^{1/n-1}$$

and zero otherwise.

THEOREM: Let X be a continuous random variable with density f_X. Suppose $g(x)$ is strictly monotone and differentiable. Let $Y=g(X)$. Then

$$\boxed{f_Y(y)=\begin{cases} f_X(g^{-1}(y))\left|\dfrac{d}{dy}g^{-1}(y)\right|, & \text{if } y=g(x) \text{ for some } x\\[2ex] 0, & \text{if } y\neq g(x) \text{ for all } x\end{cases}}$$

Proof: Suppose $g'(x)>0$ (the proof in the other case is similar). Suppose $y=g(x)$ (i.e. it is in the range of Y), then

$$F_Y(y)=P\{g(X)\le y\}=P\{X\le g^{-1}(y)\}=F_X(g^{-1}(y))$$

Differentiating both sides, we get

$$f_Y(y)=f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y)\ \blacksquare$$