Chapter 3: Continuous Random Variables

So far, we've focused on random variables suitable for counts. The range of the random variables has been $\mathbb{N}$ or $\mathbb{Z}^+$.

We may also want to model phenomena taking on values in $\mathbb{R}$, for example time to completion, time to failure, inter-arrival time, and generally values relating to duration.

Def. A random variable X on $(S, \mathcal{F}, P)$ is a function $X: S \to \mathbb{R}$ such that

$$\{s \mid X(s) \leq x\} \in \mathcal{F} \quad \text{for all } x \in \mathbb{R}.$$

Ex. If $S = \mathbb{R}$ and $\mathcal{F}$ is a σ-field containing all open intervals (a, b), then any piecewise continuous function $X: S \to \mathbb{R}$ is a valid random variable on $(S, \mathcal{F}, P)$

Def. The <u>cumulative distribution function</u>, <u>CDF</u>, F_X, associated with the random variable X is the function

$$F_X : \mathbb{R} \to [0, 1] \quad \text{satisfying}$$

$$F_X(x) = P(X \leq x) \qquad -\infty < x < \infty$$

(Note that this is the same as the definition applied to discrete random variables.)

Def. A <u>continuous random variable</u> X is a random variable for which the CDF is continuous.

We will not show this, but it is true that, for a continuous random variable, F_X will be differentiable 'almost everywhere', ie the set on which $\frac{dF}{dx}$ is undefined, say D has

$$P(X \in D) = 0.$$

Def. The <u>probability density function</u>, <u>pdf</u> of a continuous random variable X is defined by

$$f(x) = \frac{dF}{dx}(x).$$

Note $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$

$$\int_{-\infty}^{\infty} f(t)\,dt = 1$$

$$0 \leq F(x) \leq 1$$

$\lim_{x\to -\infty} F(x) = 0 \qquad \lim_{x\to\infty} F(x) = 1$

$$P(X=c) = \int_c^c f(t)\,dt = 0$$

F is monotone increasing, ie $a > b \Rightarrow$ $F(a) \geq F(b)$.

$f(x) \geq 0$ where defined.

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

for all continuous random variables X with CDF F and pdf f.

Conversely, if f is a measurable function mapping $\mathbb{R}$ to the non-negative reals satisfying

$$\int_{-\infty}^{\infty} f(t)\,dt = 1$$

then f and F defined by $F(x) = \int_{-\infty}^{x} f(t)\,dt$ are the pdf and CDF of a continuous random variable. In particular, take $S = \mathbb{R}$, $\mathcal{F}$ = the smallest σ-field containing all open intervals, and define P by $P(B) = \int_{t\in B} f(t)\,dt$. The random

variable $X: S = \mathbb{R} \to \mathbb{R}$ by
$X(s) = s$ has f and F as its
pdf and CDF, respectively.

Ex. Let X be a random variable that takes on values in $[0, \frac{1}{2}]$ only, and takes those on with equal probability. The form of the pdf f must be

$$f(x) = \begin{cases} c & x \in [0, \frac{1}{2}] \\ 0 & \text{otherwise} \end{cases}$$

(why?)

To pin this down, we know

$$1 = \int_{-\infty}^{\infty} f(t)\,dt = \int_0^{1/2} c\,dt = c\,t\Big|_0^{1/2} = \tfrac{1}{2}c,$$

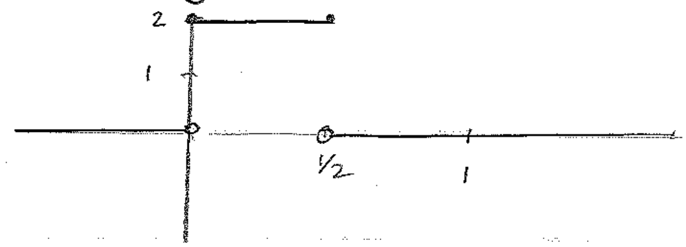
so $c = 2$.

The corresponding CDF F is

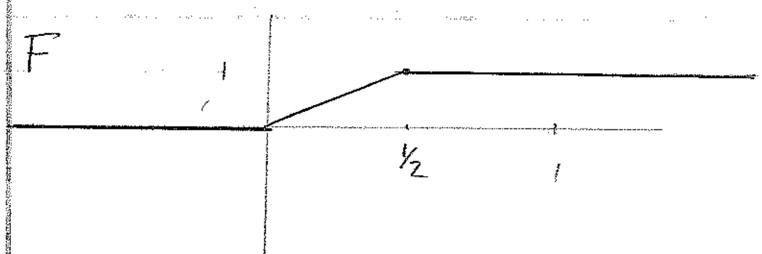
$$F(x) = \begin{cases} 0 & x \leq 0 \\ 2x & 0 \leq x \leq \frac{1}{2} \\ 1 & x \geq \frac{1}{2} \end{cases}$$

Graphically,

f:

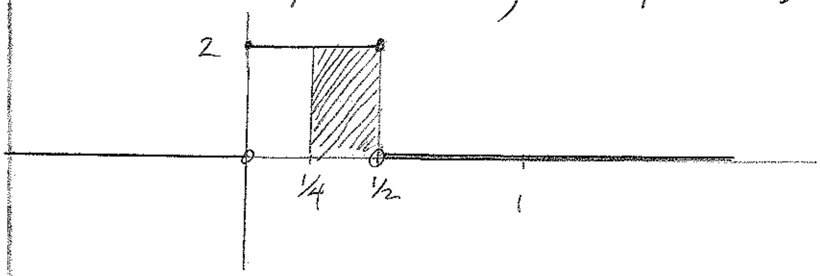


F



Recall from integral calculus that $P(a \le X \le b) = F(b) - F(a) = \int_a^b f(t)\,dt =$ area under f between a and b.

In the example above, $P(\frac{1}{4} \le X \le 1) =$



$= 2\left[\frac{1}{2} - \frac{1}{4}\right]$

$= \frac{1}{2}$

Can you think of a duration that could be modeled with a random variable like this one?

Most random variables we consider will be continuous or discrete.

The only remaining possibility is that F is continuous except on a countable set on which it has jump discontinuities. (The proof of this is outside the scope of this course.)

Ex. Model the time to failure of a component that has an immediately fatal manufacturing flaw with probability .1. If the component does not have this flaw, then the probability that the component lasts until time x drops off exponentially with x, ie

$$1 - F(x) \propto e^{-\lambda x} \quad \text{for} \quad x \geq 0 :$$

$$F(x) = 0 \qquad x < 0 \qquad \text{(easy)}$$

$$F(0) = 0.1 \qquad \text{(given)}$$

$$F(x) = 1 - ce^{-\lambda x} \qquad \text{(given)}$$

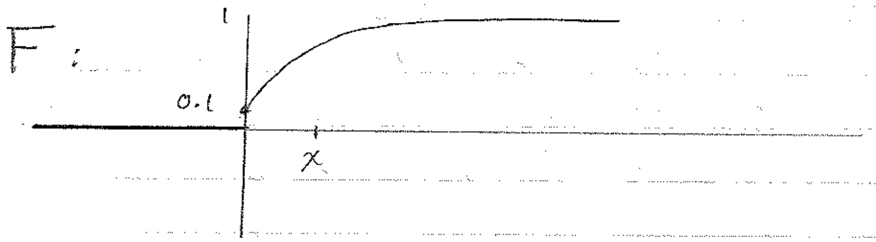
The only discontinuity is at $x = 0$, and $\lim_{x \to 0^+} F(x) = .1$, so $c = .9$

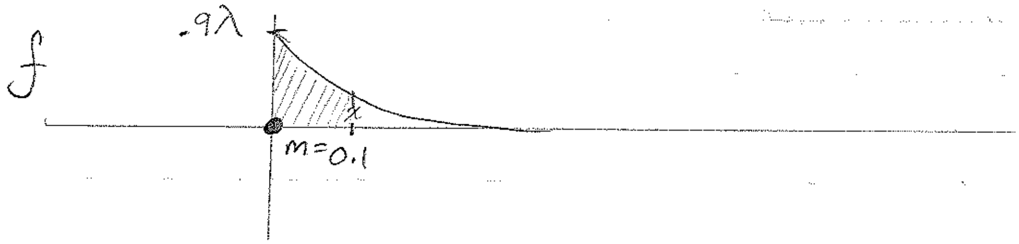
and

$$F(x) = \begin{cases} 0 & x < 0 \\ .1 & x = 0 \\ 1 - .9e^{-\lambda x} & x \geq 0 \end{cases}$$

$$f(x) = \begin{cases} 0 & x < 0 \\ \text{undefined} & x = 0 \\ .9\lambda e^{-\lambda x} & x > 0 \end{cases}$$

We say that f has a point mass of mass 0.1 at $x = 0$. The value of λ would have to be determined experimentally.





$$F(x) = \int_0^x .9\lambda e^{-\lambda t}\, dt + 0.1$$

$$\left(= -.9e^{-\lambda t} \Big|_0^x + .1 \right.$$

$$\left. = .9 - .9e^{-\lambda x} + .1 = 1 - .9e^{-\lambda x} \checkmark \right)$$

Try #1 §3.1 p.119.

3.2 The Exponential Distribution

The exponential distribution is often used to model interarrival times of jobs, service time of jobs, time to failure, and time to repair. The memoryless property makes separable networks analytically manageable.

Exponential functions are a 1 parameter family of continuous random variables.

The CDF of the exponential distribution with parameter λ is $(\lambda > 0)$

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & 0 \le x \le \infty \\ 0 & \text{otherwise} \end{cases}$$

with pdf

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

A random variable with this CDF and pdf is said to have the $EXP(\lambda)$ distribution, $X \sim EXP(\lambda)$.

The memoryless or Markov property of $EXP(\lambda)$ is stated analytically as

$X \sim EXP(\lambda)$ $t \geq 0$, $y \geq 0$ implies

$$\dagger \quad P(X \leq y+t \mid X > t) = P(X \leq y).$$

Pf. $$\frac{(1-e^{-\lambda(y+t)})-(1-e^{-\lambda t})}{1-(1-e^{-\lambda t})}$$

$$= \frac{\cancel{e^{-\lambda(t)}}(1-e^{-\lambda y})}{\cancel{e^{-\lambda t}}}$$

$$= 1-e^{-\lambda y}$$

In fact, any nonnegative continuous random variable that is memoryless, ie satisfies $\dagger$, is exponential. The proof is in the text.

Why is a distribution satisfying $\dagger$ called memoryless?

What are the implications of modeling time to failure with a memoryless distribution?

9 Important Relationship between Exponential RV and Poisson RV:

Interarrival times of Poisson RV's are exponentially distributed.

To gain intuition into why this is true, consider the time X to the first arrival if the number of arrivals in $(0, t]$, N_t, is Poisson with parameter λt.

$$P(X > t) = P(N_t = 0) = e^{-\lambda t}\frac{(\lambda t)^0}{0!}$$

$$= e^{-\lambda t}, \text{ so}$$

$$F(x) = 1 - e^{-\lambda x}$$

One can show that the numbers of arrivals over disjoint intervals are independent Poisson random variables, so this argument generalizes to interarrival times.

Ex. Jobs arrive at a CPU at an average rate of $\lambda = .05$ jobs/sec. Assuming that the number of arrivals in any time interval is Poisson distributed, what is the distribution of interarrival times? What is the probability of an interarrival time of $\leq .25$ sec?

$$X \sim EXP(.05), \text{ so}$$

$$F_X(.25) = 1 - e^{-(.05)(.25)} \approx .012$$

Try #1 §3.2 p. 123

1. Suppose CPU demand of jobs can be modeled as $EXP(\frac{1}{140 \text{ ms}})$. A job not completing in 100 ms is requeued. What is the probability that an arriving job will be requeued? Of 800 jobs, how many are expected to finish in the first quantum (100ms)?

3.3 The Reliability, Failure Density, and Hazard Function

Let X be the lifetime or time to failure of a component. The probability that the component functions at least until time t, $P(X > t) = 1 - F_X(t)$, is called the <u>reliability</u>, <u>$R(t)$</u>

Compare the probability that the component will fail in $(t, t+\Delta t]$, approximately $f(t)\Delta t$, to the conditional probability that the component fails in $(t, t+\Delta t)$ given that it functioned up to time t:

$$\frac{F(t+\Delta t) - F(t)}{1 - F(t)}$$

Note

$$\lim_{x \to 0} \frac{1}{x}\left(\frac{F(t+x) - F(t)}{1 - F(t)}\right) = \lim_{x \to 0} \frac{F(t+x) - F(t)}{x} \frac{1}{R(t)}$$

$$= \frac{f(t)}{R(t)}$$

Def. The <u>instantaneous failure rate</u>, <u>$h(t)$</u> =

$$\frac{f(t)}{R(t)}$$

This is also called the <u>hazard rate</u>, force of mortality, conditional failure rate, or failure rate.

Thus $h(t)\Delta t$ is approximately the probability that the component will fail in $(t, t+\Delta t]$ given that the component functions until time t.

Exercise: Compare $f(t)$ and $h(t)$ for $X(t)$ representing the lifetime of a human.

Ex. For $F \sim EXP(\lambda)$, $h(t) = \dfrac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$

If you're given h, you can reconstruct R.

$$R(t) = \exp\left[-\int_0^t h(x)\,dx\right]$$

Verification: $\displaystyle\int_0^t h(x)\,dx = \int_0^t \frac{f(x)}{R(x)}\,dx$

$= -\displaystyle\int_0^t \frac{R'(x)}{R(x)}\,dx$, change of variables $R = R(x)$ $dR = R'(x)dx$,

$= -\displaystyle\int_{R(0)}^{R(t)} \frac{dR}{R} = -\ln(R)\Big|_{R(0)}^{R(t)} = -\ln(R(t))$

so $R(t) = \exp\left[-\int_0^t h(x)\,dx\right]$.

Def. The cumulative hazard, $H(t)$, is defined by

$$H(t) = \int_0^t h(x)\,dx$$

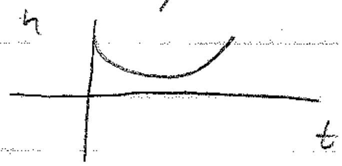
Thus $R(t) = e^{-H(t)}$

Note $h(t)$ constant $\Leftrightarrow$ $F(t)$ exponential

Exercise: Will a component that deteriorates as it ages have an increasing or decreasing hazard rate?

Can you think of a circumstance in which a component's conditional reliability improves with age?

What sort of situation is described by a bathtub shape for h?



Try problem 1, p. 129 §3.3

3.4 Some important distributions

Hypoexponential Distribution: We will see that a sum of exponential random variables with distinct parameters $\lambda_1, \lambda_2, \dots \lambda_r$ has a hypoexponential distribution. This type of distribution has been shown to be a useful model for service times for IO operations.

To express $F(t)$ and $f(t)$ set

$$a_i = \prod_{\substack{j=1 \\ j \neq i}}^{r} \frac{\lambda_j}{\lambda_j - \lambda_i}$$

$\left(\sum_{i=1}^{r} a_i = 1\right.$, though we won't prove this$\left.\right)$

Then

$$f(t) = \sum_{i=1}^{r} a_i \lambda_i e^{-\lambda_i t}$$ and

$$F(t) = 1 - \sum_{i=1}^{r} a_i e^{-\lambda_i t}$$ (add to p. 778)

For $k=2$,

$$F(t) = 1 - \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} + \frac{\lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_2 t}$$

$$f(t) = \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} \left(e^{-\lambda_1 t} - e^{-\lambda_2 t}\right)$$

$$h(t) = \frac{\lambda_1 \lambda_2 \left(e^{-\lambda_1 t} - e^{-\lambda_2 t}\right)}{\lambda_2 e^{-\lambda_1 t} - \lambda_1 e^{-\lambda_2 t}}$$

$h(0) = 0$, h is increasing, $\lim_{t \to \infty} h(t) = \min \{\lambda_1, \lambda_2\}$

Erlang Distribution: We will later show this distribution is distribution of a sum of iid exponential distributions

The pdf of the r-stage Erlang distribution with parameter λ is

$$f(t) = \frac{\lambda^r t^{r-1} e^{-\lambda t}}{(r-1)!} \qquad r \in \mathbb{N}^+$$

$$F(t) = 1 - \sum_{k=0}^{r-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \qquad (p. 778)$$

(obtained by repeated integration by parts)

~~Exercise: What random variable is a 1-stage Erlang RV?~~

The r-stage Erlang distribution models the time until the r^{th} occurrence of phenomena Poisson distributed with parameter λ.

Gamma Distribution:

$$f(t) = \frac{\lambda^\alpha t^{\alpha-1} e^{-\lambda t}}{\Gamma(\alpha)} \qquad t > 0 \quad \alpha > 0$$

where $\Gamma(\alpha)$ is the constant needed to make f a pdf:

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt$$

Note $\Gamma(1) = \int_0^\infty e^{-t}\,dt = 1$ and for $\alpha > 1$

$$\int_0^\infty t^{\alpha-1} e^{-t}\,dt = -t^{\alpha-1} e^{-t}\Big|_0^\infty$$
$$+\int_0^\infty (\alpha-1)\, t^{\alpha-2} e^{-t}\,dt = (\alpha-1)\int_0^\infty t^{\alpha-2} e^{-t}\,dt.$$

Thus $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$.

Conclude $\Gamma(n) = (n-1)!$ $\quad n \in \mathbb{Z}^+$

Also $\Gamma(1/2) = \sqrt{\pi}$, so $\Gamma(3/2) = \frac{1}{2}\sqrt{\pi}$,
$\Gamma(5/2) = \frac{3}{2}\cdot\frac{1}{2}\sqrt{\pi}$, etc.

Gam$(1/2, n/2)$ is the χ^2 distribution with n degrees of freedom, useful in statistical analyses.

Hyperexponential Distribution: motivation — you have k possible processes, each exponentially distributed with parameter λ_i, $i \in \{1, 2, \ldots k\}$, and selected with probability α_i, $\sum_{i=1}^{k} \alpha_i = 1$

$$f(t) = \sum_{i=1}^{k} \alpha_i \lambda_i e^{-\lambda_i t}$$

$$F(t) = \sum_{i=1}^{k} \alpha_i (1 - e^{-\lambda_i t})$$

Weibull: generalizes the exponential family of distributions. It offers good flexibility on the shape of the failure rate function.

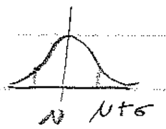
$$f(t) = \lambda\alpha\, t^{\alpha-1} e^{-\lambda t^{\alpha}}$$

$$F(t) = 1 - e^{-\lambda t^{\alpha}}$$

$$h(t) = \lambda\alpha t^{\alpha-1}$$

Log-Logistic and Pareto distributions also described.

Normal or Gaussian Distribution



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

$-\infty \leq x \leq \infty$
$\sigma, \mu \in \mathbb{R}$
$\sigma > 0$

Write $X \sim N(\mu, \sigma^2)$

$F(x)$ has no closed form. The standard normal table on p. 789 is for $Z \sim N(0,1)$ and shows

$$F_Z(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx$$

(Note the typo's in the left-hand column. They should proceed in steps of .1: 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7...)

By symmetry, $F_Z(-z) = 1 - F_Z(z)$.

For $X \sim N(\mu, \sigma^2)$, $\frac{X-\mu}{\sigma} \sim N(0,1)$

$F_X(x) = F_Z\left(\frac{x-\mu}{\sigma}\right)$ as follows:

$$\int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) dt =$$

(setting $v = \left(\frac{t-\mu}{\sigma}\right)$ $dv = \frac{1}{\sigma} dt$
$t = x \Rightarrow v = \frac{x-\mu}{\sigma}$)

$$\int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}(v)^2\right) dv =$$

$$F_Z\left(\frac{x-\mu}{\sigma}\right).$$

ie: $P\left(\frac{X-\mu}{\sigma} < t\right) = P(X < \sigma t + \mu)$

$= P\left(Z < \frac{(\sigma t + \mu) - \mu}{\sigma}\right) = P(Z < t)$

Ex. A variety of battery is manufactured by a process that makes batteries with voltages normally distributed with $\mu = 1.5\,V$ and $\sigma = .075\,V$. What is the probability that a particular battery actually delivers $1.6\,V$ or more?

$$1-F_Z\left(\frac{1.6-1.5}{.075}\right)=1-F_Z(1.\bar{3})$$

$$\approx 1-F_Z(1.33)\approx 1-.9082=.0918$$

using the normal table.

(The error function, erf, is available in some math packages.

$$\mathrm{erf}(u)=\frac{2}{\sqrt{\pi}}\int_0^u e^{-y^2}\,dy$$

This is useful for Normal calculations because, tracking down the integrals,

$$\tfrac{1}{2}\left(1+\mathrm{erf}\left(\tfrac{z}{\sqrt{2}}\right)\right)=F_Z(z)$$

In R, the CDF of $N(a,b)$ is, pnorm(x,a,b).

For intuition about N, note

$$P(-1\le Z\le 1)=2\left(F_Z(1)-\tfrac{1}{2}\right)=2F_Z(1)-1$$

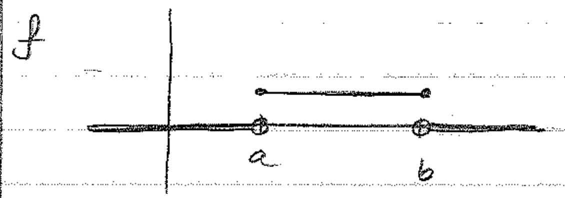
$$\approx .6826$$

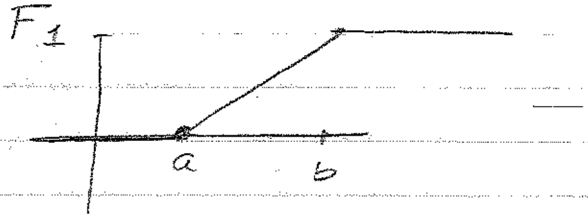
$P(-2\le Z\le 2)\approx .9642$
$P(-3\le Z\le 3)\approx .9974$
$P(-4\le Z\le 4)\approx 1.0000$

Uniform or Rectangular Distribution:

$$f(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x \le 1 \\ 1 & x \ge b \end{cases}$$





Def. A <u>defective distribution</u> X is one for which

$$\lim_{t \to \infty} F_X(t) < 1$$

Try p. 147, #1 life times $\sim N(5 \times 10^6 h, (5 \times 10^5 h)^2)$

Do you estimate that at least 95% have lifetimes greater than $4 \times 10^6 h$?

Functions of a Random Variable

Given an RV X and a function Φ : image of $X \to \mathbb{R}$,

$$Y = \Phi(X) \text{ will again be}$$

a random variable under fairly general conditions on Φ, for example piecewise continuity.

You can often get $F_Y(y)$ from F_X using inverse images:

$$F_Y(y) = P(Y \leq y) = P(\{x : \Phi(x) \leq y\})$$

Ex. $\Phi(X) = X^2$ $F_Y(y) =$

$P(\{x \mid x^2 \leq y\}) = P(-\sqrt{y} \leq X \leq \sqrt{y}) =$

$F_X(\sqrt{y}) - F_X(-\sqrt{y})$.

Differentiate to get $f_Y(y)$

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} (f_X(\sqrt{y}) + f_X(-\sqrt{y})) & y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Week 10 Proofread # 3

Example 3.9:

In particular, if $X \sim N(0,1)$ and $Y = X^2$ then

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}} \, y^{-1/2} e^{-y/2} & y > 0 \\ 0 & \text{otherwise} \end{cases}$$

Check:

This makes Y a Gamma distribution with parameters $1/2$ and $1/2$, ie χ^2_1.

Theorem: Let X be a continuous random variable with density f_X. Define the set $I \subseteq \mathbb{R}$ by

$$I = \{x \in \mathbb{R} \mid f_X(x) > 0\} \quad (\text{support of } f_X)$$

If Φ is a differentiable monotone function $\Phi : I \to \mathbb{R}$ and $Y = \Phi(X)$ then

$$f_Y(y) = \begin{cases} f_X(\Phi^{-1}(y)) \, |(\Phi^{-1})'(y)| & y \in \Phi(I) \\ 0 & \text{otherwise.} \end{cases}$$

Pf. for Φ decreasing, ie $y_0 \le y \Rightarrow \Phi(y_0) \ge \Phi(y)$

$$F_Y(y) = P(Y \le y) = P(\Phi(X) \le y)$$

$$= P(X \ge \Phi^{-1}(y)) = 1 - F(\Phi^{-1}(y))$$

thus $f_Y(y) = -f_X(\Phi^{-1}(y))(\Phi^{-1})'(y)$

$$= f_X(\Phi^{-1}(y))\,|(\Phi^{-1})'(y)|$$

An important consequence of this theorem is a method to simulate many RV's given a uniform distribution U on $[0,1]$

Given U uniform on $[0,1]$ and an RV X which has F_X strictly monotone on I, define $\Phi : [0,1] \to \mathbb{R}$ by

$\Phi = F_X^{-1}$. Set $Y = \Phi(U)$. Then

Y and X are identically distributed.

Pf. By the previous theorem,

$$f_Y(y) = \begin{cases} f_U[(F_X^{-1})^{-1}(y)]\,((F_X^{-1})^{-1})'(y) & \text{for } y \in F_X^{-1}([0,1]) \\ 0 & \text{otherwise} \end{cases}$$

ie $f_Y(y) = \begin{cases} f_U(F_X(y)) f_X(y) & y \in I_X \\ 0 & \text{otherwise} \end{cases}$

Ex. $Y = aX + b$ $\Phi^{-1}(y) = \frac{y-b}{a}$,

$(\Phi^{-1})'(y) = \frac{1}{a}$ so

$f_Y(y) = f_X\left(\frac{y-b}{a}\right) \frac{1}{|a|}$

Ex $X \sim \text{EXP}(\lambda)$, $Y = rX$, $f_Y(y) =$

$\frac{1}{r} \lambda e^{-\lambda y/r}$, so $Y \sim \text{EXP}(\lambda/r)$

Ex. Given a uniformly distributed RV U on $[0,1]$, what function of U is exponentially distributed with parameter λ.

Set $X \sim \text{EXP}(\lambda)$. $F_X(x) = 1 - e^{-\lambda x}$ $x > 0$

Calculate $F_X^{-1}(u)$

$u = 1 - e^{-\lambda x}$ $e^{-\lambda x} = 1 - u$

$-\lambda x = \ln(1-u)$, so $F_X^{-1}(u) = \frac{\ln(1-u)}{-\lambda}$.

Conclude $\frac{\ln(1-U)}{-\lambda} \sim \text{EXP}(\lambda)$

try p. 152 #3

X uniform on $[-\frac{\pi}{2}, \frac{\pi}{2}]$, $Y = \sin(X)$.
Calculate $f_Y(y)$

BTW $\arcsin'(y)$, $-\frac{\pi}{2} \le \arcsin(y) \le \frac{\pi}{2}$,

is $\frac{1}{\sqrt{1-y^2}}$, as can be seen

by applying the differentiation rule

$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$ and $\sin^2\theta + \cos^2\theta = 1$

3.6 Jointly Distributed Random Variables

Recall that if X, Y are RV's mapping $S \to \mathbb{R}$ then

$$F_{XY}(x,y) = P(X \le x \text{ and } Y \le y)$$

As in the discrete case,

$$0 \le F(x,y) \le 1 \qquad x, y \in \mathbb{R}$$

$F(x,y)$ is monotone increasing in x and y

$$\lim_{x \to -\infty} F(x,y) = \lim_{y \to -\infty} F(x,y) = 0$$

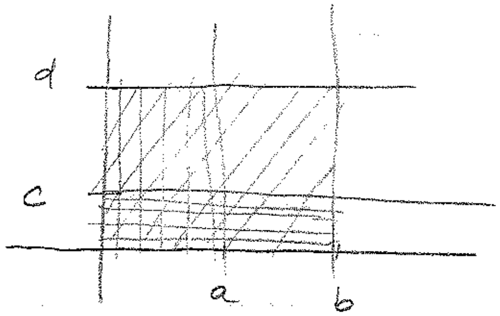
$$\lim_{x,y \to \infty} F(x,y) = 1$$

Note $F(x,y)$ is continuous if X and Y are continuous.

As before

$$P(a < X \le b \text{ and } c < Y \le d) =$$

$$F(b,d) - F(a,d) - F(b,c) + F(a,c)$$



The marginal distribution of X is

$$F_X(x) = \lim_{y \to \infty} F_{XY}(x,y),$$

likewise for Y.

Def. X and Y are <u>jointly continuous</u> if there exists a measurable function $f(x,y)$: $\mathbb{R}^2 \to \mathbb{R}$ satisfying

$$F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,dv\,du$$

Def. X and Y are independent random variables if

$$F_{XY}(x,y) = F_X(x)F_Y(y) \qquad x,y \in \mathbb{R}$$

If X and Y are continuous, this is equivalent to

$$f_{XY}(x,y) = f_X(x)f_Y(y) \qquad \text{a.e. } x,y \in \mathbb{R}$$

Example: Suppose the joint pdf of X and Y is

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) \qquad |\rho| \leq 1$$

Find the marginal density f_Y.

Rewrite $f(x,y)$ as

$$\frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\left[\frac{x^2-2\rho xy+\rho^2y^2}{2(1-\rho^2)} + \frac{(1-\rho^2)y^2}{2(1-\rho^2)}\right]\right)$$

So $\int_{-\infty}^{\infty} f(x,y)\,dx$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-y^2}{2}\right) \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(x-\rho y)^2}{2(1-\rho^2)}\right) dx}_{\text{this is the pdf for } N(\rho y, \sqrt{1-\rho^2}) \text{, so equals 1}}$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-y^2}{2}\right)$$

Thus $Y \sim N(0,1)$

Similarly, $X \sim N(0,1)$

Note X and Y are not independent.

try #3 p. 157 : $f(x,y) = \begin{cases} \frac{1}{200} & (x,y) \in A \\ 0 & \text{elsewhere} \end{cases}$

where A is the triangular region with vertices $(100,100)$, $(100,120)$, and $(120,120)$. X and Y represent the lifetimes of two series connected components. Find the reliability expression for the system.

3.7 Order Statistics

Let $X_1, X_2, \ldots X_n$ be iid RV (this is called a random sample of X with $X_i \sim X$).

Define $Y_1, Y_2 \ldots Y_n$ to be a permutation of $X_1, X_2, \ldots X_n$ satisfying

$$Y_1 \leq Y_2 \leq \ldots Y_n$$

Then Y_k is the <u>k^{th} order statistic</u> of $X_1, X_2, \ldots X_n$

Theorem: Let $X_1, X_2, \ldots X_n$ be a random sample with CDF F for each X_i. Then the CDF of Y_k is

$$F_{Y_k}(y) = \sum_{j=k}^{n} \binom{n}{j} (F(y))^j (1-F(y))^{n-j}$$

because $Y_k \le y \Leftrightarrow$ at least k of the X_i's lie in $(-\infty, y]$ (the X_i's that permute to $Y_1, \ldots Y_k$)

so

$$F_{Y_k}(y) = \sum_{j=k}^{n} P(\text{exactly } j \text{ of the } X_i\text{'s lie in } (-\infty, y])$$

$$= \sum_{j=k}^{n} \binom{n}{j} (F(y))^j (1-F(y))^{n-j}$$

In particular, $F_{min}(y) = F_1(y) = 1-(1-F(y))^n$

$\{ \quad = 1 - \binom{n}{0} F(y)^0 (1-F(y))^n =$ by the binomial theorem applied to $(F(y)+1-F(y))^n$

$$\left(\sum_{j=0}^{n} \binom{n}{j} F(y)^j (1-F(y))^{n-j} \right) - \binom{n}{0} F(y)^0 (1-F(y))^n$$

$$= \sum_{j=1}^{n} \binom{n}{j} F(y)^j (1-F(y))^{n-j} \quad \}$$

$$F_{max}(y) = F_n(y) = (F(y))^n$$

Note that if F is the CDF of the time to failure for each of n independent components,

$$R_{series}(t) = 1 - F_{min}(t) = 1 - \left(1 - [1 - F(t)]^n\right)$$
$$= (R(t))^n$$

$$R_{parallel}(t) = 1 - F_{max}(t) = 1 - (F(t))^n$$
$$= 1 - [1 - R(t)]^n$$

In general, if R_i is the reliability of the i^{th} independent component,

$$R_{series}(t) = \prod_{i=1}^{n} R_i(t)$$

$$R_{parallel}(t) = 1 - \prod_{i=1}^{n} (1 - R_i(t))$$

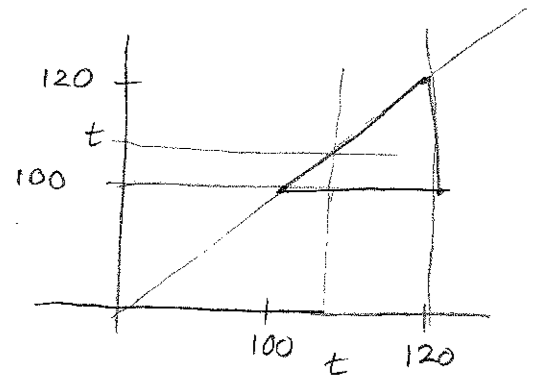
Example: Suppose the system of n components is parallel and $F(t) = 1 - e^{-\lambda t}$, $R(t) = e^{-\lambda t}$

$$F_{sys}(t) = (1 - e^{-\lambda t})^n$$

$$f_{sys}(t) = \lambda n (1 - e^{-\lambda t})^{n-1} e^{-\lambda t}$$

$$h_{sys}(t) = \frac{f_{sys}(t)}{R_{sys}(t)} = \frac{\lambda n (1 - e^{-\lambda t})^{n-1} e^{-\lambda t}}{1 - (1 - e^{-\lambda t})^n}$$

Note $h_{sys}(t)$ depends on t, though the individual h's are constant.



The probability that the system functions until time t is $P(X > t \text{ and } Y > t)$

For $t < 100$ $\quad P(X > t \text{ and } Y > t) = 1$

For $t < 120$ $\quad P(X > t \text{ and } Y > t) = 1 - \frac{1}{200} \cdot \frac{(t-100)^2}{2}$

For $t > 120$ $\quad P(X > t \text{ and } Y > t) = 0.$

Try problem 1a, p. 165: A parallel system has 3 independent components with times to failure distributed as $EXP(\lambda_i)$ $\lambda_i = .0001, .0002, .0004$

a. Determine the probability that the system will work for 1000 h.

3.8. Distribution of Sums

Suppose X and Y are jointly continuous random variables with joint density $f(x,y)$. Suppose the random variable Z is a function of X and Y, $Z = \Phi(X,Y)$. For $z \in \mathbb{R}$, define

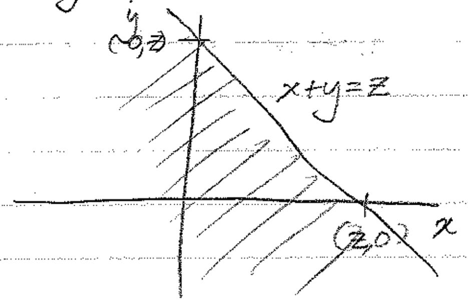
$$A_z = \{(x,y) \mid \Phi(x,y) \leq z\} = \Phi^{-1}((-\infty, z])$$

Then the CDF for Z has the formula

$$F_Z(z) = P(Z \leq z) = \iint_{A_z} f(x,y)\,dx\,dy.$$

The sum $Z = X + Y$ is a frequently used random variable.

For $Z = X + Y$, the set A_z is graphed here



So $F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f(x,y)\,dy\,dx$. Change variables by

$y = t - x \quad t = y + x$ to get

$$F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z} f(x, t-x)\,dt\,dx =$$

$$\int_{-\infty}^{z} \int_{-\infty}^{\infty} f(x, t-x)\,dx\,dt$$

Thus $f_Z(z) = \int_{-\infty}^{\infty} f(x, z-x)\,dx$

If X and Y are independent, this becomes

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,dx \qquad -\infty < z < \infty$$

$$= \int_{-\infty}^{\infty} f_Y(y) f_X(z-y)\,dy$$

Example: (3.25) Cold Spare: Two statistically identical components are available for a system. Only one is required for the system to function properly. Keep one component as a powered-off spare. Assume it does not fail in this state. Assume failure detection and switching are perfect. Run the system until the first component fails, then swap in the second. The lifetime Z of the system is $X+Y$ where X and Y are the lifetimes of the components. Now assume X and Y are iid exponentials: $f_X(x) = \lambda e^{-\lambda x}$ $x \geq 0$, $f_Y(y) = \lambda e^{-\lambda y}$ $y \geq 0$, so

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,dx \qquad z \geq 0$$

$$= \int_0^z \lambda e^{-\lambda x} \lambda e^{-\lambda(z-x)} = \int_0^z \lambda^2 e^{-\lambda z}\,dx$$

$$\lambda^2 e^{-\lambda z} x \Big|_0^z = \lambda^2 z e^{-\lambda z} \qquad z \geq 0$$

Thus Z has a 2-stage Erlang distribution.

$$F(t) = 1 - e^{-\lambda t} - \lambda t e^{-\lambda t} \qquad t \geq 0$$
$$R(t) = (1+\lambda t) e^{-\lambda t}$$

This is a special case of

Thm. If X_1, X_2, X_r are iid, $X_i \sim EXP(\lambda)$ $\forall i$, then $X_1 + X_2 + \dots X_r$ has the r-stage Erlang distribution with parameter λ. (Pf. in Ch. 4)

Ex. If X_1, X_2 independent, $X_1 \sim EXP(\lambda_1)$, $X_2 \sim EXP(\lambda_2)$ then $Z = X_1 + X_2$ has a 2-stage hypoexponential distribution $Z \sim HYPO(\lambda_1, \lambda_2)$.

Calculation: $f_Z(z) = \int_0^z \lambda_1 e^{-\lambda_1 x} \lambda_2 e^{-\lambda_2 (z-x)} dx$

$$= \lambda_1 \lambda_2 e^{-\lambda_2 z} \int_0^z e^{(\lambda_2 - \lambda_1)x} dx$$

$$= \lambda_1 \lambda_2 e^{-\lambda_2 z} \left. \frac{e^{(\lambda_2 - \lambda_1)x}}{\lambda_2 - \lambda_1} \right|_0^z$$

$$= \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_2 z} \left(e^{\lambda_2 z - \lambda_1 z} - 1 \right)$$

$$= \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 z} - \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_2 z}$$

$$= \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 z} + \frac{\lambda_1 \lambda_2}{\lambda_1 - \lambda_2} e^{-\lambda_2 z}$$

Again, this is a special case of a theorem we will not prove here.

Theorem: If $Z = \sum_{i=1}^{n} X_i$ where $X_1, X_2, \ldots X_n$ are mutually independent, $X_i \sim EXP(\lambda_i)$, and the λ_i's are distinct,

$$f_Z(z) = \sum_{i=1}^{n} a_i \lambda_i e^{-\lambda_i z} \qquad z \geq 0$$

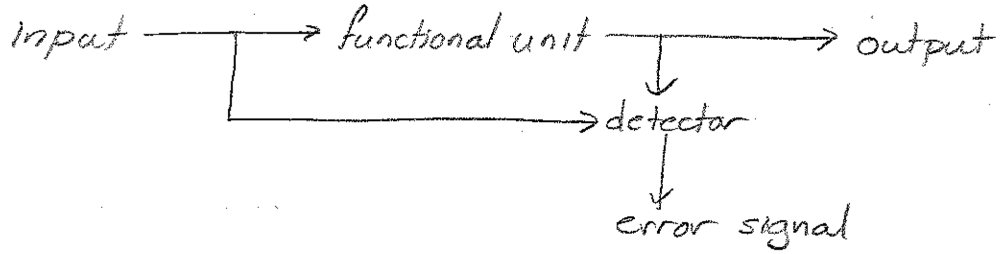
where $\quad a_i = \prod_{\substack{j=1 \\ j \neq i}}^{r} \dfrac{\lambda_j}{\lambda_j - \lambda_i}$

$(Z \sim HYPO(\lambda_1, \lambda_2, \ldots \lambda_r))$

Corollary: the sum Z of two hypoexponential distributions $X_1 \sim HYPO(\lambda_1, \lambda_2, \ldots \lambda_k)$, $X_2 \sim HYPO(\lambda_{k+1}, \ldots \lambda_r)$, X_1, X_2 independent and $\lambda_1, \lambda_2, \ldots \lambda_r$ distinct has

$Z \sim HYPO(\lambda_1, \lambda_2, \ldots \lambda_r)$ (Why?)

Example 3.29



T: time to failure of unit
C: time to failure of detector
D: time to detection of unit failure
(Detector failure is reported instantaneously.)

X: time to indication of failure of functional unit or detector

Y: time to first failure occurrence (functional unit or detector)

$X = \min\{C, T+D\}$
$Y = \min\{T, C\}$

Suppose T, C, and D are independent exponential distributions with parameters λ_T, λ_C, and λ_D respectively.

For Y, $F_Y(y) = 1-(1-F_T(y))(1-F_C(y))$
$= 1-e^{-(\lambda_T+\lambda_C)y}$
so $Y \sim EXP(\lambda_T+\lambda_C)$

For X, $T+D$ is hypoexponentially distributed

$$F_{T+D}(t) = 1-\frac{\lambda_T}{\lambda_T-\lambda_D}e^{-\lambda_D t}-\frac{\lambda_D}{\lambda_D-\lambda_T}e^{-\lambda_T t}$$

so the apparent reliability is

$$R_X(t) = P(X \geq t) = P(\min\{T+D, C\} \geq t)$$

$$= P(T+D \geq t)P(C \geq t)$$
$$= \left(\frac{\lambda_T}{\lambda_T-\lambda_D}e^{-\lambda_D t}+\frac{\lambda_D}{\lambda_D-\lambda_T}e^{-\lambda_T t}\right)e^{-\lambda_C t}$$

Consider a system with n iid components with failure law $\sim EXP(\lambda)$. If k or more components must be functioning for the system to function then the lifetime of the system is given by the order statistic Y_{n-k+1}.

($n-k+1$ components have failed $\rightarrow n-(n-k+1) = k-1$ are working, ie the system just failed)

Theorem 3.5 If $X_1, X_2, \ldots X_n$ are iid with $X_i \sim EXP(\lambda)$ then

$$Y_{n-k+1} \sim HYPO[n\lambda, (n-1)\lambda, \ldots k\lambda],$$

an $n-k+1$ stage hypo exponential distribution.

Pf. Proceed by induction on $n-k+1$.

base case: $n-k+1=1$, ie $n=k$

$Y_1 = \min(X_1, X_2, \ldots X_n) \sim EXP(n\lambda)$,

as required.

Assume $Y_{n-j+1} \sim HYPO(n\lambda, (n-1)\lambda, \cdots j\lambda)$.

Show $Y_{n-(j-1)+1} = Y_{n-j+2} \sim$
$HYPO(n\lambda, (n-1)\lambda, \ldots j\lambda, (j-1)\lambda)$
given the assumption:

$Y_{n-j+2} = Y_{n-j+1} + \min\{W_{n-j+2}, \ldots W_n\}$

where the W_i's are the lifetimes of the surviving components. By the memoryless property, $W_i \sim EXP(\lambda)$.
There are $j-1$ W_i's, so the minimum is distributed as $EXP((j-1)\lambda)$, a 1 stage hypo exponential distribution.

We have previously shown that a sum of hypoexponential distributions with distinct parameters is hypoexponentially distributed with the set of parameters equal to the union of the parameters of the summands.
Conclude $Y_{n-j+2} \sim HYPO(n\lambda, (n-1)\lambda, \ldots j\lambda, (j-1)\lambda)$ ✓

A simplification of the coefficients appears on p. 179.

The following diagram shows another way to think of this result:

$$\underset{\substack{\uparrow \\ Y_1}}{EXP(n\lambda)} + \underset{\substack{\uparrow \\ Y_2}}{EXP((n-1)\lambda)} + \cdots \underset{\substack{\uparrow \\ Y_{n-k+1}}}{EXP(k\lambda)} + \quad \cdots + \underset{\substack{\uparrow \\ Y(n)}}{EXP(\lambda)}$$

Example 3.32 p.179 similarly applies induction and the memoryless property of exponential distributions to compute the lifetime of a k-out-of-n system with m spares.

Suppose we have a k out of n system with n components and m spares. Active components' lifetimes are iid $EXP(\lambda)$, while inactive components independently have time to failure $\sim EXP(\mu)$. Then system lifetime, $L(k|n,m)$, satisfies

$$L(k|n,m) \sim HYPO(n\lambda+\mu, n\lambda+2\mu, \ldots n\lambda+m\mu, k\lambda, (k+1)\lambda, \ldots n\lambda)$$

The corresponding diagram,

$EXP(n\lambda+m\mu) \to EXP(n\lambda+(m-1)\mu) \to \cdots EXP(n\lambda+\mu) \to EXP(n\lambda) \to$

$EXP((n-1)\lambda) \to \cdots EXP(k\lambda)$

reflects the fact that, while there are p spares, the next component failure occurs at the minimum of n $EXP(\lambda)$ RV's and p $EXP(\mu)$ RVs. The system then has n active components and $p-1$ spares.

Once all the spares have been exhausted, the system is k out of n system, previously analyzed.

Try problem 1, p. 182: Suppose you have a source of n random values, $\mu_1, \dots \mu_n$, uniformly distributed on $[0,1]$. Give a formula in the μ_i's that results in a random variable that is hypoexponentially distributed with parameters $\lambda_1, \lambda_2, \dots \lambda_n$.

(Hints: If U is uniform on $[0,1]$,
$F_X^{-1}(U) \sim X$.
If you had $X_i \sim EXP(\lambda_i)$ $i=1,\dots n$, how could you generate a $Z \sim HYPO(\lambda_1, \lambda_2, \dots \lambda_n)$?)

3.9 Functions of Normal Random Variables

Normal distributions are common approximating distributions in statistics due to the Central Limit Theorem. As a result, distributions of functions of Normal RV's acquire importance. Some results are given below.

Theorem 3.6 $X_1, X_2, \ldots X_n$ mutually independent $X_i \sim N(\mu_i, \sigma_i^2)$ then $S_n = \sum_{i=1}^{n} X_i$

satisfies
$$S_n \sim N(\Sigma \mu_i, \Sigma \sigma_i^2)$$

Proposition: $X \sim N(\mu, \sigma^2)$ then $aX \sim N(a\mu, a^2\sigma^2)$

Can you show this based on $F_{aX}(t) = F_X(\frac{t}{a})$

so $f_{aX}(x) = \frac{1}{a} f(\frac{x}{a})$?

Example 3.34: Suppose $X_1, X_2, \ldots X_n$ iid, $X_i \sim N(\mu, \sigma^2)$. Set $\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n}$.

Then $\bar{X} \sim N\left(\sum_{i=1}^{n} \frac{\mu}{n}, \sum_{i=1}^{n} \frac{\sigma^2}{n^2}\right) = N\left(\mu, \frac{\sigma^2}{n}\right)$

The distribution of the mean $\bar{X}$ has the same center as the X_i's and a sharper peak. This is why the mean of normally distributed measurements centered on some value μ is a better estimate of μ than a single measurement.

Sums of Squares of iid $X_1, \ldots X_n$, $X_i \sim N(0,1)$

If $X \sim N(0,1)$ then $Y = X^2$ satisfies

$$Y \sim GAM(1/2, 1/2) = \chi_1^2$$

$$(\chi_n^2 = GAM(1/2, \tfrac{n}{2})$$

We showed this as Example 3.9

Example 3.35: If X_1, X_2 iid, $X_i \sim N(0,1)$ then $Y = X_1^2 + X_2^2$ satisfies

$$Y \sim EXP(1/2) = GAM(1/2, 1) = \chi_2^2$$

This is proved by hand in the text.

More generally

Theorem 3.7: If $X_1, X_2 \cdots X_n$ iid standard normal and $Y = \sum_{i=1}^{n} X_i^2$ implies

$$Y \sim GAM(1/2, n/2) = \chi^2_n$$

This is a consequence of

Theorem 3.8: Let $X_1, \ldots X_n$ be a sequence of mutually independent random variables with

$$X_i \sim GAM(\lambda, \alpha_i)$$

Set $S_n = \sum_{i=1}^{n} X_i$. Then

$$S_n \sim GAM(\lambda, \sum_{i=1}^{n} \alpha_i)$$

Application: (glimpse of Ch. 10) If you know $X_1, \ldots X_n$ iid, $X_i \sim N(\mu, \sigma^2)$ with μ known and σ^2 unknown and you want to approximate σ^2 from values of the X_i's (aka 'data'), note

$\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2_n$. If we use

$\frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^2 = S^2$, say, as an estimator

for σ^2, we can obtain a confidence interval for our result.

For example, with $n=10$, and a 95% confidence interval required, use the χ^2 table to find values a and b such that $X \sim \chi^2_{10}$ implies

$$P(a < X < b) = .9.$$

If we pick $a = 3.24697$ and $b = 20.4832$

$$P(a \le \chi^2_{10}) = .975 \qquad \text{and}$$

$$P(\chi^2_{10} \le b) = .025 \qquad \text{so}$$

$$P(a \le \chi^2_{10} \le b) = .95$$

Then

$$P\left(a \le \frac{10S^2}{\sigma^2} \le b\right) = .95$$

$$P\left(\frac{1}{b} \le \frac{\sigma^2}{10S^2} \le \frac{1}{a}\right) = .95$$

$$P\left(\frac{10S^2}{b} \le \sigma^2 \le \frac{10S^2}{a}\right) = .95,$$

so

$$P(.49S^2 \le \sigma^2 \le 3.08S^2) = .95$$

$$P(.7S \le \sigma \le 1.8S) = .95$$

Now we are prepared to say what we mean by a 95% confidence interval: If we calculate the upper and lower bounds from the RVs X_i, the probability that the interval will include the value of σ^2 is .95

Actually, often μ and σ^2 are unknown. Then the estimator

$$S^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}$$ is used.

It turns out that $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$,

making the calculations for the confidence intervals for this estimator very similar to those above.

Theorem 3.9 Let Y_1 and Y_2 be independent RVs $Y_1 \sim \chi^2_{n_1}$, $Y_2 \sim \chi^2_{n_2}$ then Z

$= \frac{Y_1/n_1}{Y_2/n_2}$ has the F_{n_1, n_2} distribution,

for which tables are available.
Pf in text.

Theorem 3.10 If V and W are independent random variables such that $V \sim N(0,1)$ and $W \sim \chi^2_n$ then

$$T = \frac{V}{\sqrt{W/n}}$$ has the t-distribution

with n degrees of freedom.

The pdf of the t-distribution with n degrees of freedom is

$$f_{T_n}(t) = \frac{\Gamma(n+1/2)}{\sqrt{n\pi}\,\Gamma(n/2)} \left[1 + \frac{t^2}{n}\right]^{-(n+1)/2}$$

As $n \to \infty$, $f_{T_n}(t)$ approaches $f_Z(t)$, the pdf of the standard normal random variable.

Application: This gives us an estimator for μ of a random sample $X_1, X_2, \ldots X_n$, $X_i \sim N(\mu, \sigma^2)$ if both μ and σ are unknown.

Recall $V = \dfrac{(\bar{X}-\mu)\sqrt{n}}{\sigma} \sim N(0,1)$. Also

$$W = \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n} \left[\frac{X_i - \bar{X}}{\sigma}\right]^2 \sim \chi^2_{n-1}.$$

Thus $$T = \frac{V}{\sqrt{\frac{W}{n-1}}} = \frac{(\bar{X}-\mu)\sqrt{n}}{\sigma\sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}}} = \frac{\bar{X}-\mu}{S/\sqrt{n}}$$

has the t-distribution with $n-1$ degrees of freedom.

To find a confidence interval for μ, use a t-table to find a such that $1 - F_{t_n}(a) = \alpha/2$, say. Then, by the symmetry of the t distribution around 0

$$P(-a \le T \le a) = 1-\alpha$$

$$P\left(-a \le \frac{\bar{X}-\mu}{S/\sqrt{n}} \le a\right) = 1-\alpha$$

$$P\left(-\frac{aS}{\sqrt{n}} \le \bar{X}-\mu \le \frac{aS}{\sqrt{n}}\right) = 1-\alpha$$

$$P\left(-\bar{X}-\frac{aS}{\sqrt{n}} \le -\mu \le -\bar{X}+\frac{aS}{\sqrt{n}}\right) = 1-\alpha$$

$$P\left(\bar{X}-\frac{aS}{\sqrt{n}} \le \mu \le \bar{X}+\frac{aS}{\sqrt{n}}\right) = 1-\alpha$$

Try #3 §3.9 p.190.