Chapter 4: Expectation

Def. The expectation $E[X]$ of a random variable is given by

$$E[X] = \begin{cases} \sum_i x_i\, p(x_i) & X \text{ discrete} \\ \int_{-\infty}^{\infty} x f(x)\,dx & X \text{ continuous} \end{cases}$$

and, in general $\int_S x f(x)\,dx$ using a generalized definition of integration

if $\sum_i |x_i| p(x_i) < \infty$, $\int_{-\infty}^{\infty} |x| f(x)\,dx < \infty$, $\int_S |x| f(x)\,dx < \infty$ respectively.

The expectation, aka expected value of an RV is a weighted mean with more likely values weighted more heavily.

Ex. If X is the number of available channels in a system with five channels, each of whose availability is a Bernoulli trial with $p = 1/2$, what is the expectation of X, ie the expected number of available channels?

Recall $X \sim B(5, 1/2)$

$$E(X) = \sum_{k=0}^{5} k \binom{5}{k} \frac{1}{32} =$$

$$0 \cdot \frac{1}{32} + 1 \frac{5}{32} + 2 \frac{10}{32} + 3 \frac{10}{32} + 4 \cdot \frac{5}{32} + 5 \frac{1}{32}$$

$$= \frac{80}{32} = \frac{5}{2}.$$

(The expected value is not necessarily a value X takes on with nonzero probability.)

Expected value of geometric RV:

$$E(X) = \sum_{i=1}^{\infty} i p (1-p)^{i-1} = p \sum_{i=1}^{\infty} i (1-p)^{i-1}$$

$$= p \frac{d}{dp} \sum_{i=0}^{\infty} -(1-p)^{i} = p \frac{d}{dp} \left(\frac{-1}{1-(1-p)} \right) = p \frac{1}{p^2} = \frac{1}{p}$$

Expected value of uniform distribution:

$P(X=i) = \frac{1}{n} \quad i \in \{1, 2, .. n\} \qquad E[X] = \sum_{i=1}^{n} \frac{i}{n}$

$$= \frac{1}{n} \sum_{i=1}^{n} i = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}$$

Expected value of X distributed according to Zipf's law (Ex. 4.2):

Zipf's Law $\quad p(i) = \frac{c}{i} \quad i = 1, 2, \ldots n,$

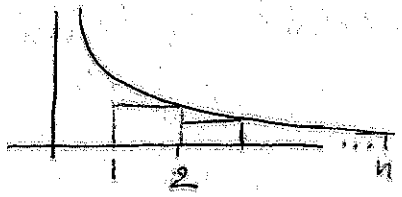
$$c = \left(\sum_{i=1}^{n} \frac{1}{i} \right)^{-1} = \frac{1}{H_n}$$

Application: a list is ordered so that the first item is one requested with probability $\frac{c}{1}$, 2^{nd} with probability $\frac{c}{2}$, i^{th} with probability $\frac{c}{i}$. In this case $E[X]$ is the expected number of items examined in a sequential search for a requested item.

$$E[X] = \sum_{i=1}^{n} i\frac{c}{i} = \sum_{i=1}^{n} c = \frac{n}{H_n}$$

Note H_n is bounded as follows

$$\ln(n) < H_n < \ln(n)+1.$$



so $$\frac{n}{\ln(n)+1} < E[X] < \frac{n}{\ln(n)}$$

Expected Value of Exponential Distribution:

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

$$E[X] = \int_0^\infty \lambda x e^{-\lambda x} dx \qquad u = \lambda x \quad dv = e^{-\lambda x}$$

$$v = -\frac{e^{-\lambda x}}{\lambda}$$

$$= -xe^{-\lambda x}\Big|_0^\infty + \int_0^\infty e^{-\lambda x} dx$$

$$= -\frac{e^{-\lambda x}}{\lambda}\Big|_0^\infty = 0 - \left(\frac{-e^0}{\lambda}\right) = \frac{1}{\lambda}$$

try #4 §4.1 p. 197

$$F(x) = \begin{cases} 0 & x < 0 \\ p_0 & x = 0 \\ p_0 + (1-p_0)(1-e^{-\lambda x}) & x > 0 \end{cases}$$

4.2 Moments

If X is a random variable and Y, $Y = \phi(X)$ is also a random variable, then $E[Y]$ may be defined.

By definition, $E[Y] = \sum_i y_i \Pr(y_i)$ or $\int y f_Y(y)\,dy$.
In fact,

$$E[Y] = \sum_i \phi(x_i)\, p_X(x_i) \quad \text{or}$$

$$\int \phi(x) f_X(x)\,dx$$

if the relevant quantity is absolutely convergent.

This is easy to understand from an example.

Ex. A particular user is downloading information with probability .1, uploading with probability .05 and neither with probability .85
Say $X(\text{downloading}) = -1$, $X(\text{uploading}) = 1$, $X(\text{neither}) = 0$ and $Y = X^2$

$E[X] = (-1)(.1) + 0(.85) + 1(.05) = -.05$

$E[Y] = (-1)^2(.1) + 0^2(.85) + 1^2(.05) = .15$

(compare $P(Y=1) = .15$, $P(Y=0) = .85$

$E[Y] = 1(.15) + 0(.85) = .15$

Note it is NOT true that $E[\phi(X)] = \phi[E(X)]$ — here $(-.05)2 = .025 \neq .15$

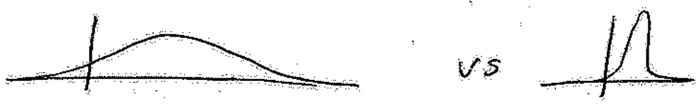
Def. The $\underline{k^{th} \text{ moment}}$ of X is $E[X^k]$.
The $\underline{k^{th} \text{ central moment}}$ of X, μ_k is $E[(X-E(X))^k]$.

Thm. If X and Y are random variables and $E[X^k] = E[Y^k]$ for all $k = 1, 2, \ldots$ then X and Y have the same distribution.

Def. The $\underline{\text{variance}}$ of a random variable X is

$$Var[X] = \mu_2 = \sigma_X^2 = \begin{cases} \sum_i (x_i - E(x))^2 p(x_i) & X \text{ discrete} \\ \int_{-\infty}^{\infty} (x - E(x))^2 f(x)\,dx & X \text{ contin.} \end{cases}$$

Note the variance will be larger when the distribution has a higher probability of values distant from the mean



Ex. Recall $\int_0^\infty x^{\alpha-1} e^{-\lambda x}\,dx = \frac{\Gamma(\alpha)}{\lambda^\alpha}$ $\left(= \frac{(\alpha-1)!}{\lambda^\alpha} \text{ if } \alpha \in \mathbb{N}\right)$.
Calculate σ^2 for $X \sim EXP(\lambda)$

$$\int_0^\infty \left(x - \frac{1}{\lambda}\right)^2 \lambda e^{-\lambda x}\,dx =$$

$$\int_0^\infty \left(x^2 - \frac{2x}{\lambda} + \left(\frac{1}{\lambda}\right)^2\right) \lambda e^{-\lambda x}\,dx$$

$$= \lambda\left(\frac{\Gamma(3)}{\lambda^3} - \frac{2\Gamma(2)}{\lambda^3} + \frac{\Gamma(1)}{\lambda^3}\right)$$

$$= \frac{1}{\lambda^2}(2 - 2 + 1) = \frac{1}{\lambda^2}$$

Note If $Y = aX + b$, $E[Y] = aE[X] + b$.

Try # 4 § 4.2 p. 200

4.3 Expectation Based on Multiple Random Variables

If $X_1, \ldots X_n$ random vector and $Y = \phi[X_1, \ldots X_n]$ then

$$E[Y] = E[\phi(X_1, \ldots X_n)] = \begin{cases} \sum_{\vec{x}_i} \phi(\vec{x}_i)\, p(\vec{x}_i) & \text{discrete} \\ \int \cdots \int_{\mathbb{R}^n} \phi(\vec{x}) f(\vec{x})\, dx_1 \ldots dx_n & \text{contin.} \end{cases}$$

Ex 4.7 Consider a moving head on a cylinder with outer radius b and inner radius a. Approximate track radius by a continuous value. Let X be the start position of the head. Assume X is uniform on $[a,b]$. Let Y be the end position, assumed uniform on $[a,b]$. The seek distance is $|X-Y|$. Find $E[|X-Y|]$.

$$\frac{1}{(b-a)^2} \int_a^b \int_a^b |x-y|\, dx\, dy = \iint_{a \le y \le x \le b} (x-y)\, dx\, dy + \iint_{a \le x \le y \le b} (y-x)\, dx\, dy$$

$$= \frac{2}{(b-a)^2} \int_a^b \int_a^x (x-y)\, dy\, dx = \frac{2}{(b-a)^2} \int_a^b x(x-a) - \frac{(x^2-a^2)}{2}\, dx$$

$$= \frac{2}{(b-a)^2} \int_a^b \frac{1}{2}x^2 - ax + \frac{a^2}{2}\, dx = \frac{2}{(b-a)^2}\left(\frac{1}{6}(b^3-a^3) - \frac{a}{2}(b^2-a^2) + \frac{a^2}{2}(b-a) \right)$$

$$= \frac{1}{3(b-a)^2}\left[b^3 - a^3 - 3ab^2 + \cancel{3a^3} + 3a^2 b - \cancel{3a^3} \right]$$

$$= \frac{1}{3(b-a)^2}(b-a)^3 = \frac{b-a}{3}$$

(In practice, generally smaller.)

Theorem 4.1 $E[X+Y] = E[X] + E[Y]$
(Note X and Y need not be independent.)

pf. for discrete case (continuous case is in the text)

$$E[X+Y] =$$

$$\sum_{x_i}\sum_{y_j} (x_i + y_j)\, p(x_i, y_j) =$$

$$\sum_{x_i}\sum_{y_j} x_i p(x_i, y_j) + \sum_{y_j}\sum_{x_i} y_j\, p(x_i, y_j)$$

$$= \sum_{x_i} x_i \Big(\sum_{y_j} p(x_i, y_j)\Big) + \sum_{y_j} y_j \sum_{x_i} p(x_i, y_j)$$

$$= \sum_{x_i} x_i\, p(x_i) + \sum_{y_j} y_j\, p(y_j)$$

$$= E[X] + E[Y]$$

Note this makes the expectation of a binomial RV easy. Thm. 4.1 generalizes to $E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i]$. The expected value of a Bernoulli trial is $1 \cdot p + 0(1-p) = p$. $X \sim B(n,p)$ is the sum of n Bernoulli trials, so $E[X] = np$.

Recall that the expectation of a geometric random variable is $\frac{1}{p}$. A negative Binomial RV with parameters p and r is the sum of r geometric RV's with parameter p. $E[\text{negBin}(r,p)] = \frac{r}{p}$.

Ex. What is the expected number of children if a couple has children until having 2 sons? Assuming that the sexes of the children are independent and each is male or female with $p = \frac{1}{2}$, get $\frac{2}{.5} = 4$

Cor. $\sigma^2 = E[X^2] - (E[X])^2$

$$E[(X - E[X])^2] = E[X^2 - 2X\,E[X] + (E[X])^2]$$

$$= E[X^2] - 2E[X]E[X] + (E[X])^2$$
$$= E[X^2] - (E[X])^2$$

Cor. $Y = aX + b \Rightarrow \sigma_Y^2 = a^2 \sigma_X^2$

Thm. 4.2 If X and Y are independent

$E[XY] = E[X]E[Y]$

(The proof is another reordering proof, using $p(x_i, y_j) = p(x_i)p(y_j)$)

Note $E[X]E[Y] = E[XY]$ does not imply independence.

Theorem 4.3 If X and Y are independent RVs

$Var[X+Y] = Var[X] + Var[Y]$

Pf. $Var[X+Y] = E[(X+Y)^2] - (E[X+Y])^2$

$= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 =$

$= E[X^2] + 2\cancel{E[X]E[Y]} + E[Y^2] - ((E[X])^2 + 2\cancel{E[X]E[Y]} + (E[Y])^2)$

$= E[X^2] - (E[X])^2 + E[Y^2] - (E[Y])^2$

Ex. $X_1, \dots X_n$ iid $\rightarrow Var(\sum_i \frac{X_i}{n}) = \sum \frac{1}{n^2} Var[X_i]$

$= \frac{1}{n}\sigma^2$ $\qquad (\sigma^2 = Var[X_i] \;\; i = 1, 2 \dots n)$

Ex. The variance of a Bernoulli trial is $p(1-p)$ so the variance of $X \sim B(n,p)$ is $np(1-p)$.

Def. X, Y random variables,

$$E[(X-E[X])(Y-E[Y])] =$$

$$E[XY]-E[X]E[Y] = \underline{Cov[X,Y]}$$ the

<u>covariance</u> of X and Y.

Note $Cov[X,Y]=0$ if X, Y independent, but not conversely.

Ex. If $Y = aX+b$

$$Cov[X,Y] = E[X(aX+b)] - E[X](aE[X]+b)$$

$$= aE[X^2] + bE[X] - (a(E[X])^2 + bE[X])$$

$$= a\, Var[X] = \frac{1}{a} Var[Y]$$, so

$$Cov^2[X,Y] = Var[X]Var[Y] \text{ if } Y = aX+b.$$

In general $0 \leq Cov^2[X,Y] \leq Var[X]Var[Y]$

with $Cov^2[X,Y] = Var[X]Var[Y] \Leftrightarrow Y = aX+b$.

pf omitted.

Def. The <u>correlation coefficient</u> of X and Y

$$\rho(X,Y) = \frac{Cov[X,Y]}{\sqrt{Var[X]\,Var[Y]}}$$

$$-1 \le \rho(X,Y) \le 1$$

with $\rho(X,Y) = \begin{cases} -1 & \text{if } X = -aY+b \quad a>0 \\ 0 & X,Y \text{ uncorrelated} \\ 1 & \text{if } X = aY+b \quad a>0 \end{cases}$

Try #1, 4 pp. 206-207 § 4.3

4.4 Transform Methods

As we saw with the PGF, sometimes manipulating the transform of a random variable is easier than the corresponding manipulation of the variable itself.

Further, we can get mean, variance, and higher moments directly from certain transforms. This will allow us to extract summary data for distributions we're unable/unwilling to calculate barehanded.

Def The <u>Moment Generating Function</u> of a random variable X, $M(\theta)$ is defined by

$$M(\theta) = \begin{cases} \sum_j e^{x_j\theta} p_j & \text{if } X \text{ is discrete} \\ \int e^{x\theta} f(x)\,dx & \text{if } X \text{ is continuous} \end{cases}$$

In general, the MGF of X is $E[e^{X\theta}]$, when defined. In most cases of practical interest, $M(\theta)$ will be defined for some interval of values of θ.

Note the PGF defined earlier is $M_X(\ln(z))$

$$= \sum_{j=0}^{\infty} e^{j\ln(z)} p_j = \sum_{j=0}^{\infty} p_j z^j$$

Def. The Fourier transform $N_X(\tau)$ of a random variable X is defined by

$$N_X(\tau) = M_X(i\tau) \qquad i = \sqrt{-1}$$

$$= E[e^{i\tau X}]$$

$N_X(\tau)$ is defined for all real values of τ for all random variables X.

Def. The Laplace Stieljes transform of a non-negative continuous random variable X is

$$L(s) = M(-s) = \int_0^\infty e^{-sx} f(x)\,dx.$$

(This transform is commonly used in Queuing Theory.)

Thm 4.4 Linear Translation: Let $Y = aX+b$. Then $M_Y(\theta) = e^{b\theta} M_X(a\theta)$.

Pf. $E[e^{\theta Y}] = E[e^{\theta(aX+b)}] =$
$E[e^{\theta b} e^{a\theta X}] = e^{\theta b} E[e^{(a\theta)X}] = e^{\theta b} M_X(a\theta)$

Cor. $Y = aX+b$ implies $N_Y(\tau) = e^{i\tau b} N_X(a\tau)$
$L_Y(s) = e^{-bs} L_X(as)$

Note also that $\int_{-\infty}^{\infty} e^{x\theta}(f(x)+g(x))\,dx =$

$$\int_{-\infty}^{\infty} e^{x\theta} f(x)\,dx + \int_{-\infty}^{\infty} e^{x\theta} g(x)\,dx.$$

The Convolution Theorem: Let $X_1, X_2, \ldots X_n$ be mutually independent random variables. Set $Y = \sum_{i=1}^{n} X_i$. If $M_{X_i}(\theta)$ is defined for all $i \in \{1,2,\ldots n\}$

then $M_Y(\theta)$ is defined and

$$M_Y(\theta) = \prod_{i=1}^{n} M_{X_i}(\theta)$$

$$\text{Pf. } M_Y(\theta) = E[e^{\theta Y}] = E[e^{\theta X_1 + \theta X_2 + \cdots \theta X_n}]$$

$$= E\left[\prod_{i=1}^{n} e^{\theta X_i}\right] = \underbrace{\prod_{i=1}^{n} E[e^{\theta X_i}]}_{\text{by independence}} = \prod_{i=1}^{n} M_{X_i}(\theta)$$

Correspondence Theorem: If $M_{X_1}(\theta) = M_{X_2}(\theta)\ \forall \theta$ then $F_{X_1}(x) = F_{X_2}(x)\ \forall x$, likewise for N and L.

proof is beyond the scope of this course.

Impact: If we recognise the form of the MGF of a random variable X_1 as matching that of a known distribution X_2, then the distribution of X_1 is of the same type as that of X_2.

Application: i) If $X \sim EXP(\lambda)$ then $L_X(s) = \frac{\lambda}{\lambda+s}$

Verification: $\int_0^{\infty} e^{-sx} \lambda e^{-\lambda x}\, dx =$

$$\frac{\lambda}{\lambda+s} \int_0^{\infty} (\lambda+s) e^{-(\lambda+s)x}\, dx = \frac{\lambda}{\lambda+s}\left[-e^{-(\lambda+s)x} \Big|_0^{\infty}\right] = \frac{\lambda}{\lambda+s}$$

ii. If $X \sim ERL(\lambda, r)$ $r \in \mathbb{Z}^+$, $\lambda > 0$ then
$$\mathcal{L}_X(s) = \left(\frac{\lambda}{\lambda+s}\right)^r$$

By induction on r:

$r=1$. This reduces to part i.

Assume the formula holds for r.
If $X \sim ERL(\lambda, r+1)$,
$$\mathcal{L}_X(s) = \int_0^\infty e^{-sx} \frac{\lambda e^{-\lambda x}(\lambda x)^r}{r!}\,dx$$
$$= \frac{\lambda}{\lambda+s}\int_0^\infty \underbrace{\frac{(\lambda x)^r}{r!}}_{u} \underbrace{(\lambda+s)e^{-(\lambda+s)x}}_{dv}\,dx$$
$$= \frac{\lambda}{\lambda+s}\left[\cancelto{0}{-e^{-(\lambda+s)x}\frac{(\lambda x)^r}{r!}\Big|_0^\infty} + \int_0^\infty e^{-(\lambda+s)x}\frac{r\lambda(\lambda x)^{r-1}}{r!}\,dx\right]$$
$$= \frac{\lambda}{\lambda+s}\int_0^\infty e^{-sx}\lambda e^{-\lambda x}\frac{(\lambda x)^{r-1}}{(r-1)!}\,dx = \frac{\lambda}{\lambda+s}\left(\frac{\lambda}{\lambda+s}\right)^r$$
$$= \left(\frac{\lambda}{\lambda+s}\right)^{r+1} \checkmark$$

iii. If $X_1, X_2, \ldots X_r$ are iid $EXP(\lambda)$ then
$$\sum_{i=1}^n X_i \sim ERL(\lambda, r).$$

This follows because the Laplace transforms are equal.

A similar computation, requiring partial fractions, confirms that if $X_1, \ldots X_n$ are mutually independent and $X_i \sim EXP[\lambda_i]$ and $i \neq j \Rightarrow \lambda_i \neq \lambda_j$ then

$$\sum_{i=1}^{n} X_i \sim HYPO[\lambda_1, \lambda_2, \ldots \lambda_n]$$

If $X \sim N(\mu, \sigma^2)$, a calculus exercise with a little complex analysis at the end shows

$$N_X(\tau) = \exp[\,i\tau\mu - \tau^2\sigma^2/2\,]$$

Using this, we can verify that the sum of mutually independent Gaussian RVs is again Gaussian:

Let $X_1, X_2, \ldots X_n$ be mutually independent random variables with $X_i \sim N(\mu_i, \sigma_i^2)$. Set $Y = \sum_{i=1}^{n} X_i$.

$$N_Y(\tau) = \prod_{j=1}^{n} e^{i\tau\mu_j - \tau^2\sigma_j^2/2} =$$

$$\exp\left[\,i\tau\sum_{j=1}^{n}\mu_j - \tau^2\left(\sum_{j=1}^{n}\sigma_j^2\right)/2\,\right]$$

Conclude $Y \sim N\left(\sum_{j=1}^{n}\mu_j, \sum_{j=1}^{n}\sigma_j^2\right)$.

The moment generating property of the MGF, $E[e^{X\theta}]$ is another reason for studying the MGF:

$$\frac{d^k}{d\theta^k} M_X(\theta)\Big|_0 = E[X^k],$$

$$(-1)^k\frac{d^k}{ds^k} L_X(s)\Big|_0 = E[X^k], \quad (-i)^k\frac{d^k}{d\tau^k} N_X(\tau)\Big|_{\tau=0} = E[X^k]$$

These formulas stem from the power series representation

$$e^{X\theta} = \sum_{i=0}^{\infty} \frac{(X\theta)^i}{i!}. \quad \text{It can be shown that}$$

$$M(\theta) = E\left[\sum_{i=0}^{\infty} \frac{(X\theta)^i}{i!}\right] = \sum_{i=0}^{\infty} E[X^i] \frac{\theta^i}{i!},$$

so

$$\frac{d}{d\theta} M(\theta) = \sum_{i=1}^{\infty} E[X^i] \frac{\theta^{i-1}}{(i-1)!}, \quad \frac{d}{d\theta} M(\theta)\Big|_0 = E[X].$$

$$\frac{d^k}{d\theta^k} M(\theta) = \sum_{i=k}^{\infty} E[X^i] \frac{\theta^{i-k}}{(i-k)!}, \quad \frac{d^k}{d\theta^k} M(\theta)\Big|_0 = E[X^k]$$

4.5 Moments and Transforms of Some Important Distributions

If $X \sim N(\mu, \sigma^2)$, $N_X(\tau) = e^{i\tau\mu - \tau^2\sigma^2/2}$, as noted previously. Thus

$$-i\frac{d}{d\tau} N_X(\tau) = -i\left(i\mu - 2\tau\frac{\sigma^2}{2}\right) e^{i\tau\mu - \tau^2\sigma^2/2}$$

$$-i\frac{d}{dt} N_X(0) = \mu \quad \text{so } E[X] = \mu$$

$$i^2 \frac{d^2}{d\tau^2} N_X(\tau) = -1\left(-\sigma^2 + (i\mu - \tau\sigma^2)^2\right) e^{i\tau\mu - \tau^2\sigma^2/2}$$

so $E[X^2] = \sigma^2 + \mu^2$ and

$$Var[X] = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$$

Note that the mean and variance of distributions, when defined, are included in the distribution's entry in the tables starting on p. 777

try 228 #2 a, b, c, d. Recall the definition of the coefficient of variation is $\dfrac{\sqrt{Var[X]}}{E[X]}$

$$F_X(t) = 0.6(1-e^{-10t}) + 0.4(1-e^{-t}).$$

a. calculate $f(x)$
b. calculate $E[X]$
c. calculate $Var[X]$
d. calculate the coefficient of variation.

Also, calculate the probability that $X \geq 1.5$

4.6 Mean Time to Failure

Let X be a random variable that gives the time to failure of a system. The CDF, $F(t)$, can be interpreted as the probability that the system survives to time t. $E[X]$ is the mean time to failure.

In this context, it is common to view $E[X]$ as

$$\int_0^\infty R(t)\,dt,$$

where $R(t) = 1 - F(t)$, so $R'(t) = -f(t)$.

Justification: Assuming $E[X]$ defined,

$$E[X] = \int_0^\infty t f(t)\,dt = -\int_0^\infty t R'(t)\,dt =$$

$$-tR(t)\Big|_0^\infty + \int_0^\infty R(t)\,dt.$$

But $-tR(t)\Big|_0^\infty = 0$ (because

$$\lim_{t\to\infty} tR(t) = \lim_{t\to\infty} t\int_t^\infty f(s)\,ds \le \int_t^\infty s f(s)\,ds \to 0)$$

so $E[X] = MTTF = \int_0^\infty R(t)\,dt$

MTTF for a Series System

If $R_i(t)$ is the reliability for component i, and the components are independent,

$R(t)$ = probability that all components are working at time t

$$= \prod_{i=1}^{n} R_i(t), \quad \text{so}$$

$$MTTF_{sys} = \int_0^\infty \prod_{i=1}^{n} R_i(t)\,dt$$

Working directly with the time-to-failure distribution of the system, X_{sys}, may be easier.

Ex. For a series system with mutually independent lifetime distributions $X_1, \ldots X_n$, $X_i \sim EXP(\lambda_i)$,

$R(t) = \exp\left[-\left(\sum_{i=1}^{n} \lambda_i\right) t\right]$. The lifetime of the system is again exponentially distributed with parameter $\sum_{i=1}^{n} \lambda_i$. $MTTF = \dfrac{1}{\sum_{i=1}^{n} \lambda_i}$.

(Why could you say that a series system is weaker than its weakest link?)

MTTF for a Parallel System:

$R(t)$ = probability that not all components have failed by time t

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

$$E[X_{sys}] = \int_0^\infty 1 - \prod_{i=1}^{n} (1 - R_i(t)) dt$$

Ex. Consider a parallel system with iid component lifetimes that are exponentially distributed with parameter λ.

Rather than evaluating the integral above, recall that $X_{sys} \sim HYPO(n\lambda, (n-1)\lambda, \dots \lambda)$

(thm. 3.5 p. 178)

so $E[X_{sys}] = \sum_{i=1}^{n} \frac{1}{i\lambda} = \frac{1}{\lambda} \sum_{i=1}^{n} \frac{1}{i}$.

(Note $\sum_{i=1}^{n} \frac{1}{i}$ grows slowly for large n.

Relate this to the notion of diminishing returns on having many components in parallel.

Ex. MTTF with Standby Redundancy:

Components with lifetimes $X_1, X_2, \ldots X_n$, not necessarily independent, are used in order as the previous component fails.

$$X_{sys} = \sum_{i=1}^{n} X_i$$

$$MTTF_{sys} = \sum_{i=1}^{n} MTTF_i$$

If the X_i's are mutually independent,

$$Var[X_{sys}] = \sum_{i=1}^{n} Var[X_i]$$

Ex. MTTF of a k out of n System with m Warm Spares:

If the lifetimes of active components are iid $EXP(\lambda)$ and the lifetimes of spares are iid $EXP(\mu)$, in the $L(k|n, m)$ arrangement

$$X_{sys} \sim HYPO(n\lambda+m\mu, n\lambda+(m-1)\mu, \ldots n\lambda+2\mu, n\lambda+\mu, n\lambda, (n-1)\lambda, \ldots k\lambda), \text{ so}$$

$$E[X_{sys}] = \sum_{i=1}^{m} \frac{1}{(n\lambda+i\mu)} + \sum_{i=k}^{n} \frac{1}{i\lambda}$$

Try #3 p. 235. Recall $h(t) = \frac{f(t)}{R(t)}$, $\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$,

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

4.7 Inequalities and Limit Theorems

There are a couple inequalities for fairly general distributions that depend only on the mean, or the mean and the variance.

Markov Inequality: If X is a non negative random variable and $E[X] = \mu$ defined, then

$$P(X \geq t) \leq \frac{\mu}{t}$$

pf. Set $Y = \begin{cases} 0 & X < t \\ t & X \geq t \end{cases}$.

Note $Y \leq X$ so $E[Y] \leq E[X]$

$p_Y(0) = P(X \leq t)$. $p_Y(t) = P(X \geq t)$

$$E[Y] = t\,p_Y(t) = t\,P(X \geq t) \leq E[X] = \mu$$

$$\text{so} \quad P(X \geq t) \leq \frac{\mu}{t}$$

Ex. Suppose all we know about a system is MTTF = 3 months. Find an upper bound on the probability that the system lasts 1 year

$$P(X \geq 12) \leq \frac{3}{12} = \frac{1}{4}$$

Note this is a very crude estimate. With specific distributions you will usually get a much tighter bound, ie lower value on RHS

Chebyshev Inequality: Let X be a random variable for which the mean μ and the variance σ^2 are defined. Then

$$P(|X-\mu| \geq t) \leq \frac{\sigma^2}{t^2} \qquad t>0$$

Pf. $P[(X-\mu)^2 \geq t^2] \leq \frac{E[(X-\mu)^2]}{t^2} = \frac{\sigma^2}{t^2}$

by Markov Inequality, and $(X-\mu)^2 \geq t^2 \Leftrightarrow |X-\mu| > t$

This implies

$$P(|X-\mu| \geq \tfrac{k}{\sigma}) \leq \frac{1}{k^2} \quad \text{and}$$

$$P(|X-\mu| < \tfrac{k}{\sigma}) \geq 1-\frac{1}{k^2}$$

Ex. Recall that if $X_1, \ldots X_n$ are iid with mean μ and variance σ^2, then $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ has mean μ and variance $\frac{\sigma^2}{n}$.

$$P(|\bar{X}-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \quad \text{, so}$$

$$\lim_{n\to\infty} P(|\bar{X}-\mu| \geq \varepsilon) = 0 \qquad \text{"weak law of large numbers"}$$

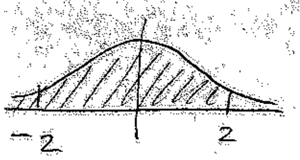
Ex. Suppose that dimensions of a type of unit vary with mean 10 cm and variance .01 cm^2. Give an upper bound on the probability that a unit will be outside the range 9.8 cm to 10.2 cm.

$$P(|X-10| \geq .2) \leq \frac{.01}{.04} = .25$$

(Compare $X \sim N(10, .01)$; $\sigma = .1$

$$P(|X-10| \geq .2)$$

$$= P\left(\left|\frac{X-10}{.1}\right| \geq 2 \right) = P(Z > 2)$$



$$= 2(1 - .9772) = .0674$$

vs )

Central Limit Theorem: Let $X_1, \ldots X_n \ldots$ be independent random variables with finite means $E[X_i] = \mu_i$ and finite variance $Var[X_i] = \sigma_i^2$ $i = 1, 2, \ldots n, \ldots$ Let Z_n be the normalized random variable

$$Z_n = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}$$

Then, under certain regularity conditions (that include $\mu_i \equiv \mu$, $\sigma_i^2 \equiv \sigma$), the limiting distribution of Z_n is standard normal in the sense that

$$\lim_{n \to \infty} F_{Z_n}(t) = P(Z_n \leq t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \, dy$$

In particular, if the X_i's are iid with mean μ and variance σ^2,

$$Z_n = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \to N(0,1)$$

Note this does not apply to all iid X_i's. If the X_i's have Cauchy distributions

$$f(x) = \frac{1}{\pi(1+x^2)}.$$

$$\int_{-\infty}^{\infty} \frac{x^2}{\pi(1+x^2)}\, dx = \infty \qquad \text{because}$$

$$\lim_{x \to \infty} \frac{x^2}{\pi(1+x^2)} = \frac{1}{\pi}.$$

Thus these X_i's do not satisfy the hypotheses for the Central Limit Theorem.

Try problem 1 p. 242