# REVIEW OF BASIC PROBABILITY AND STATISTICS

Probability and statistics are an integral part of simulation. They are needed to

1. Model a stochastic (i.e., non-deterministic) system,
2. Validate the system model,
3. Choose input probability distributions,
4. Generate random samples from these distributions,
5. Analyze the output data
6. Design simulation experiments

**Probability**: is a set function defined on a reference set S of events.

$$P(A) \equiv \text{ probability of occurrence of } A$$

(where $A$ is a subset of reference set S)

$$P(\Phi) = 0 \; ; \quad P(\mathrm{S}) = 1 \; ; \quad 0 \le P(.) \le 1$$

If $AB = \Phi$ (i.e., $A, B$ disjoint) then

$$P(A+B) = P(A) + P(B)$$

Otherwise :

$$P(A+B) = P(A) + P(B) - P(AB)$$

A **random experiment** is a procedure whose outcome is uncertain. An **event** is a possible outcome of such a random experiment. Thus, space of all possible outcomes of a random experiment can be mapped into a reference set such that each possible event becomes a subset of the reference set.

Given a random experiment whose possible outcomes are the events *A, B, C,* etc.

$P(A)$ is defined as the probability that $A$ occurs where $0 \le P(.) \le 1$

Suppose that the random experiment is repeated $N$ times and $A$ occurs $n_A$ times, then

$$P(A) = \lim_{N \to \infty} \frac{n_A}{N}$$

**Equally likely** events $A$, $B$ implies that $P(A) = P(B)$.

**Mutually exclusive** (**disjoint**) events $A$, $B$ implies $P(A+B)=P(A) + P(B)$.

If $A_1, A_2, A_3, \cdots, A_N$ constitute a set of $N$ mutually exclusive equally likely events (possible outcomes of a random experiment) then

$$P(A_1) = P(A_2) = \cdots = P(A_N) = \frac{1}{m} \;; \quad m \geq N$$

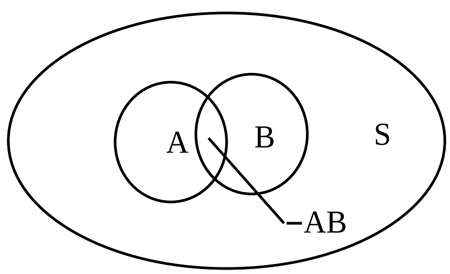
If $A_1, A_2, A_3, \cdots, A_N$ represent a mutually exclusive set of events which describe all possible outcomes of a random experiment then

$$P(A_1) + P(A_2) + \cdots + P(A_N) = 1$$

If, furthermore, the $A_i$ s are equally likely

$$P(A_1) = P(A_2) = \cdots = P(A_N) = \frac{1}{N}$$

**Conditional probability**:




**Definition**: Probability that $B$ occurs, given that $A$ does occur.

$$P(B|A) \equiv \frac{P(AB)}{P(A)} \quad \text{or} \quad P(AB) = P(A)P(B|A)$$

**Non-independence**:

$$P(ABC) = P(A)P(BC|A) = P(AB)P(C|AB) = P(A)P(B|A)P(C|AB)$$

If $P(B|A) = P(B)$ then $B$ does not depend on $A$ occurring. $A$, $B$ are said to be **independent**. Then

$$P(AB) = P(A)P(B)$$

Given non-disjoint events $A$, $B$:

$$\begin{aligned}
P\left(\overline{A}\,\overline{B}\right) &= P\left(\overline{A+B}\right) \quad \text{(see Venn diagram)} \\
&= 1 - P(A+B) \\
&= 1 - [P(A) + P(B) - P(AB)] \\
&= 1 - P(A) - P(B) + P(AB) \\
&\quad \text{Now assume events are independent (not disjoint*)} \\
&= 1 - P(A) - P(B) + P(A)P(B) \\
&= [1 - P(A)]\,[1 - P(B)] \\
&= P\left(\overline{A}\right)P\left(\overline{B}\right)
\end{aligned}$$

*In general* for independent events:

$$P\left(\overline{A_1}\,\overline{A_2}\cdots\overline{A_N}\right) = P\left(\overline{A_1 + A_2 + \cdots + A_N}\right) = P\left(\overline{A_1}\right)P\left(\overline{A_2}\right)\cdots P\left(\overline{A_N}\right)$$

*

A **random variable** (RV) represents the association of a number with the outcome of a random experiment. This is another example of a set function.

If $X$ is the outcome of a random experiment, then for each outcome there is one and only one value for $X$.

For a coin flip: Head $\Leftrightarrow$ $X = 1$ and Tail $\Leftrightarrow$ $X = 0$

Random variables may be discrete or continuous. For instance:

- numbers on a die
- direction of a spinning pointer

---

* Can a pair of disjoint events $A$ and $B$ also be independent? Explain.

**Distribution Functions (Cumulative Density Functions or cdfs):**

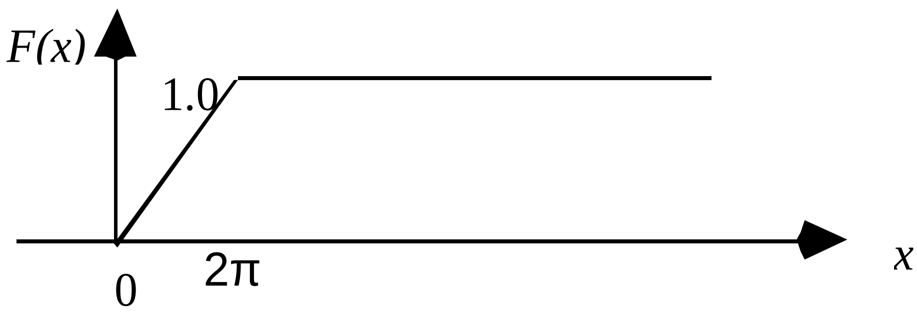
$F(x) = P(X \leq x)$ where $-\infty < x < \infty$

$F(-\infty) = 0$

$F(+\infty) = 1$

$F'(x) \geq 0$ (increasing slope)

**Example:** Spinning of a pointer $0 < X \leq 2\pi$




**Discrete Cumulative Density Function:**

$F(X) = P(X \leq x)$ (this is a continuous function)

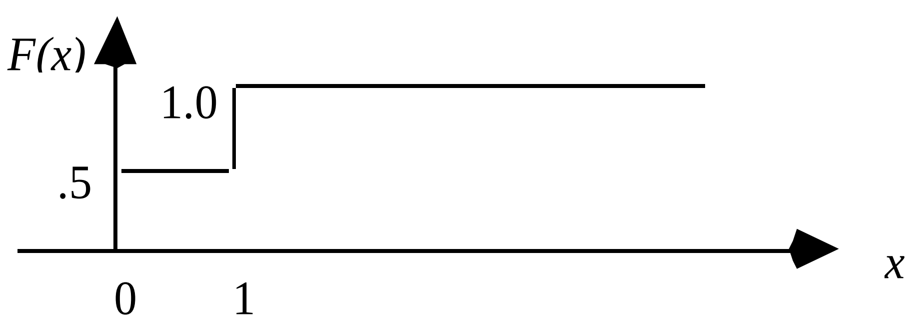
$$= \sum_{n=-\infty}^{\infty} f_n u(x - x_n)$$

where $f_n = P(X = x_n)$ and

$u(x) = 0$ for $x < 0$

$= 1$ for $x \geq 0$

**Example:** Tossing of a balanced coin where Head $\Leftrightarrow$ X = 1 and Tail $\Leftrightarrow$ X = 0



**Frequency functions:**

These are also known as **Probability Density Functions (pdfs).**

For a continuous random variable *X*:

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

$$F(x) = \int_{-\infty}^{x} f(u)du$$

$$f(x)dx = P(x - \tfrac{1}{2}dx < X \le x + \tfrac{1}{2}dx)$$

$$= F(x + \tfrac{1}{2}dx) - F(x - \tfrac{1}{2}dx)$$

$$f(x) \ge 0$$

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

*For* spinning pointer :

$$f(x) = \frac{1}{2\pi} \quad \text{for } 0 < x \le 2\pi$$

$$= 0 \quad \text{elsewhere}$$

For a discrete random variable *X*:

$$f(x) = \frac{dF(x)}{dx} = \frac{d}{dx}\left[\sum_{n=-\infty}^{\infty} f_n u(x - x_n)\right]$$

$$= \sum_{n=1}^{\infty} f_n \delta(x - x_n) \quad \text{where } \delta(.) \text{ is an impulse function}$$

$$\sum_{n=1}^{\infty} f_n = 1 \; ; \quad f_n \ge 0$$

*For* coin tossing:

$$f(x) = \tfrac{1}{2}\delta(x) + \tfrac{1}{2}\delta(x-1)$$

**Joint Random Variables:**

X and Y are joint random variables defined by:

$$F(x,y) \equiv \text{joint cumulative distribution}$$
$$\equiv P(X \le x \text{ and } Y \le y)$$

$F(x,y)$ must satisfy :

$$\mathrm{F}(-\infty,-\infty) = 0$$
$$\mathrm{F}(\infty,\infty) = 1$$
$$\frac{\partial F}{\partial x} \ge 0 \quad ; \quad \frac{\partial F}{\partial y} \ge 0$$

$$f(x,y) \equiv \text{joint PDF} \equiv \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

$(two - \text{dimensional function})$

$$= \frac{P(x - \frac{1}{2}dx < X \le x + \frac{1}{2}dx \ , y - \frac{1}{2}dy < Y \le y + \frac{1}{2}dy)}{\partial x \partial y}$$

$$\ge 0$$

*where*

$$F(\infty,\infty) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$$

**Marginal distributions:**

$$F_1(x) \equiv \text{Marginal distribution of X}$$
$$\equiv P(X \le x) = \int_{-\infty}^{x} du \int_{-\infty}^{\infty} f(u,y)dy$$
$$F_2(y) \equiv P(Y \le y) = \int_{-\infty}^{y} dv \int_{-\infty}^{\infty} f(v,y)dx$$
$$f_1(x) = \text{Marginal density function of X}$$
$$= F_1'(x) = \int_{-\infty}^{\infty} f(x,y)dy$$
$$f_2(y) = F_2'(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

**Conditional frequency functions:**

$$f_3(x \mid y) \equiv Conditional\ PDF\ of\ X\ given\ Y = y$$
$$\equiv \frac{f(x,y)}{f_2(y)}$$
$$f_4(y \mid x) \equiv Conditional\ PDF\ of\ Y\ given\ X = x$$
$$\equiv \frac{f(x,y)}{f_1(x)}$$
$$f(x,y) = f_2(y) f_3(x \mid y) = f_1(x) f_4(y \mid x)$$

If $X$, $Y$ are independent:

$$f_3(x \mid y) = f_1(x)\ and\ f_4(y \mid x) = f_2(y)$$
$$so\ that$$
$$f(x,y) = f_1(x) f_2(y)$$

*Generally, for independence*:

$$f(x_1, x_2, \quad , x_n) = f_1(x_1) f_2(x_2), \quad , f_n(x_n)$$

**Expected Values:**

| | Continuous Distribution | Discrete Distribution |
|---|---|---|
| pdf (frequency function) | *f(x)* | $\sum_n f_n \delta(x - x_n)$ |
| Expected value or mean - $E[g(X)] = \overline{g(X)} = \langle g(X) \rangle$ | $\int_{-\infty}^{\infty} g(x) f(x) dx$ | $\sum_n g(x_n) \quad f_n$ |
| Moments about zero | $\int_{-\infty}^{\infty} x^m f(x) dx$ | $\sum_n x_n^m \quad f_n$ |
| Moments about mean | $\int_{-\infty}^{\infty} (x - \mu)^m f(x) dx$ | $\sum_n (x_n - \mu)^m \quad f_n$ |

$Var[X]$ = is the variance of the distribution:

$$
\begin{aligned}
Var[X] &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx \\
&= \overline{X^2} - 2\mu\overline{X} + \mu^2 \\
&= \overline{X^2} - \mu^2 \\
&= E[X^2] - (E[X])^2
\end{aligned}
$$

$\sigma$ = the standard deviation of the distribution

$$\equiv \sqrt{Var(X)} = \sqrt{\overline{X^2} - \mu^2}$$

Other properties:

$$E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$$

Expectation of product of independent random variables:

$X, Y$ independent

$$
\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \quad f(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \quad f(x)f(y)dxdy \\
&= \int_{-\infty}^{\infty} xf(x)dx \int_{-\infty}^{\infty} yf(y)dy \\
&= E(X) \quad E(Y) = \mu_X \quad \mu_Y
\end{aligned}
$$

Let's consider **Example 5.3** in the third edition of Banks, Carson, Nelson and Nichol:

Life of a laser inspection device is a random variable $X$ given by an exponential distribution with a mean of 2 years.

**Exponential pdf:**

$$f(x) = \lambda e^{-\lambda x} \qquad 0 \le x \le \infty$$

**Mean of an exponential distribution:**

$$\mu = E[X] = \lambda \int_0^\infty x e^{-\lambda x} dx$$

Recall from calculus that

$$\int u dv = uv - \int v du$$

Letting $u = x$ and $dv = e^{-\lambda x} dx$

$$\mu = \frac{\lambda x}{-\lambda} e^{-\lambda x} \Big|_0^\infty - \frac{\lambda}{\lambda^2} e^{-\lambda x} \Big|_0^\infty$$

$$= 0 + \frac{\lambda}{\lambda^2}[1 - 0] = \frac{1}{\lambda}$$

So that for Example 6.3

$$f(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}} \qquad 0 \le x \le \infty$$

$$= \frac{1}{2} e^{-\frac{x}{2}} \qquad 0 \le x \le \infty$$

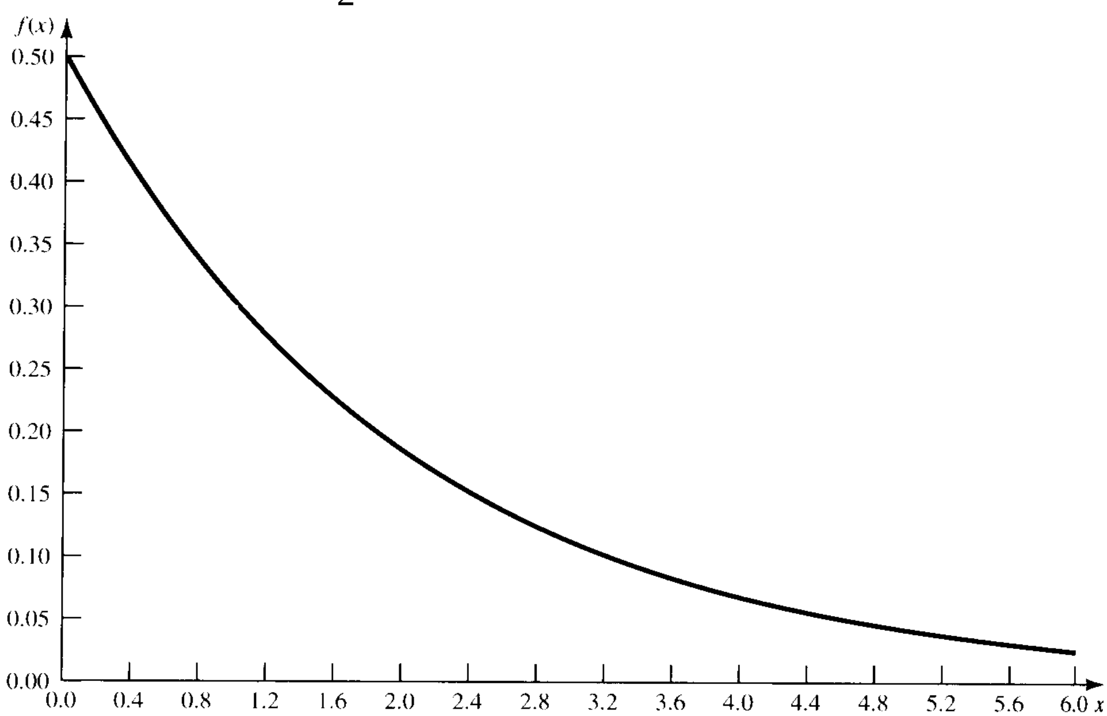



**Figure 6.3.** PDF for laser ray life.

**Exponential Distribution function:**

$$F_X(x) = \int_0^x \lambda e^{-\lambda u} du$$
$$= \frac{\lambda}{-\lambda} e^{-\lambda u}\Big|_0^x$$
$$= 1 - e^{-\lambda x} = 1 - e^{-x/2}$$

The probability that the life of the laser inspection device is between 2 and 3 years is

$$P(2 \le X \le 3) = (1 - e^{-3/2}) - (1 - e^{-2/2})$$
$$= e^{-1} - e^{-3/2}$$
$$= 0.368 - 0.223 = 0.145$$

**Variance of an exponential distribution:**

$$Var[X] = E[X^2] - (E[X])^2$$
$$E[X^2] = \lambda \int_0^\infty x^2 e^{-\lambda x} dx$$

Again using $\int u dv = uv - \int v du$

Letting $u = x^2$ and $dv = e^{-\lambda x} dx$

so that $v = -\frac{1}{\lambda} e^{-\lambda x}$ and $du = 2x dx$

$$E[X^2] = \frac{\lambda x^2}{-\lambda} e^{-\lambda x}\Big|_0^\infty + \frac{2}{\lambda} \lambda \int_0^\infty x e^{-\lambda x} dx$$
$$= 0 + \frac{2}{\lambda} E[X] = \frac{2}{\lambda^2}$$
$$(E[X])^2 = \frac{1}{\lambda^2}$$
$$Var[X] = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} = \sigma^2$$
$$\sigma = \frac{1}{\lambda}$$

**Binary distribution:**

$X$ is a discrete binary variable with values 0,1.

$$P(X=1)=p \quad \text{and} \quad P(X=0)=1-p=q$$

$$f_X(x)=q\delta(x)+p\delta(x-1)$$

$$E[X]=\mu=0\cdot q+1\cdot p=p$$

$$E[X^2]=0\cdot q+1\cdot p=p$$

$$Var[X]=p-p^2=p(1-p)=pq$$

$X$ is known as a Bernoulli trial

**Binomial distribution:**

$X_1$, $X_2$, … , $X_N$ are a sequence of independent Bernoulli trials with binary distribution parameter $p$.
Consider now the variable

$$X = X_1 + X_2 + \ldots + X_N$$

The distribution function of $X$ is just defined by the values

$$f_n = P(X=n)$$

But, this is the probability of exactly $n$ successes among the $N$ Bernoulli trials, which is given by

$$f_n = \binom{N}{n} p^n q^{N-n}$$

where $\binom{N}{n}$ are the ways in which $n$ successes can happen in $N$ trials and $p^n q^{N-n}$ is their probability.
Clearly,

$$E[X]=p+p+\cdots+p=Np$$

$$Var[X]=pq+pq+\cdots+pq=Npq$$

**Poisson distribution:** This is a limiting case of the binomial distribution.

$$N \to \infty;\ p \to 0;\ \text{such that } Np = \lambda.$$

We find the Poisson distribution as follows:

$$f_0 = \binom{N}{0} p^0 q^N$$

$$= (1-p)^N = \left(1 - \frac{\lambda}{N}\right)^N$$

$$\log f_0 = N \log\left(1 - \frac{\lambda}{N}\right)$$

Using a Taylor expansion :

$$\log(1 + a) = a - \tfrac{1}{2} a^2 + \tfrac{1}{3} a^3 - \cdots$$

so that for a small $\frac{\lambda}{N}$, $\quad \log f_0 = -\lambda - \frac{\lambda^2}{2N} - \cdots$

Since $N \to \infty$

$$f_0 \approx e^{-\lambda}$$

$$\frac{f_n}{f_{n-1}} = \frac{\dfrac{N!}{n!(N-n)!} p^n q^{N-n}}{\dfrac{N!}{(n-1)!(N-n+1)!} p^{n-1} q^{N-n+1}}$$

$$= \frac{N-n+1}{n} \frac{p}{q} \approx \frac{Np}{n} = \frac{\lambda}{n}$$

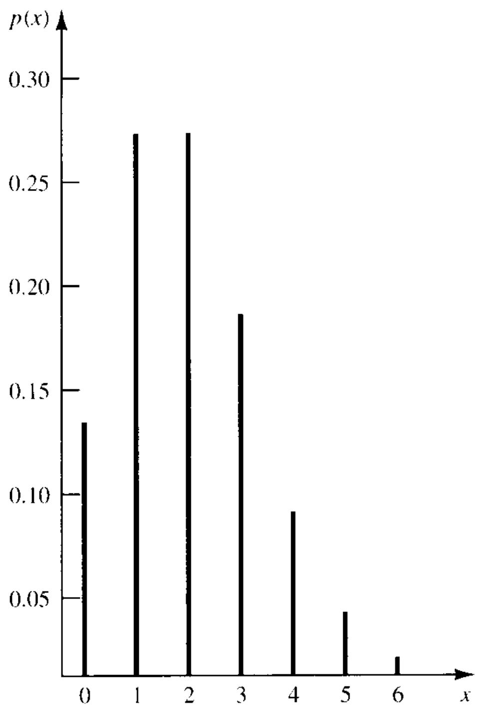
So $\quad f_1 \quad \frac{\lambda}{1} f_0 = \lambda e^{-\lambda}$

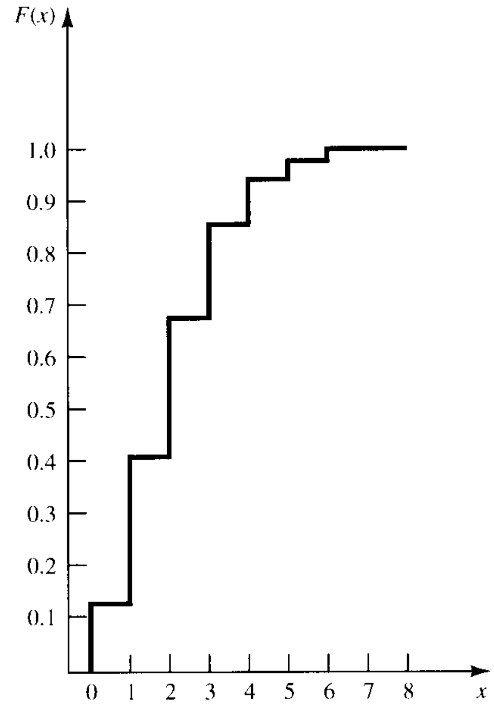
$$f_2 \quad \frac{\lambda}{2} \lambda e^{-\lambda} = \frac{\lambda^2}{2!} e^{-\lambda}$$

$$f_3 \quad \frac{\lambda}{1} \frac{\lambda^2}{2!} e^{-\lambda} = \frac{\lambda^3}{3!} e^{-\lambda}$$

$$\vdots$$

$$f_n \quad \frac{\lambda^n}{n!} e^{-\lambda}$$




**Mean and variance of the Poisson distribution:**

$$E[X] = \sum_{n=1}^{\infty} n \frac{\lambda^n}{n!} e^{-\lambda} = \lambda \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} e^{-\lambda}$$

Letting $m = n - 1$

$$E[X] = \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} e^{-\lambda} = \lambda$$

$$E[X^2] = \sum_{n=1}^{\infty} n^2 \frac{\lambda^n}{n!} e^{-\lambda} = \lambda \sum_{n=1}^{\infty} n \frac{\lambda^{n-1}}{(n-1)!} e^{-\lambda}$$

$$= \lambda \sum_{m=0}^{\infty} (m+1) \frac{\lambda^m}{m!} e^{-\lambda}$$

$$= \lambda \sum_{m=0}^{\infty} m \frac{\lambda^m}{m!} e^{-\lambda} + \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} e^{-\lambda} = \lambda \;\; \lambda + \lambda \;\; 1$$

$$= \lambda^2 + \lambda$$

So $Var[X] = \lambda^2 + \lambda - \lambda^2 = \lambda$

The cumulative Poisson distribution is tabulated in Table A.4 of Banks, Carson, Nelson and Nicol.

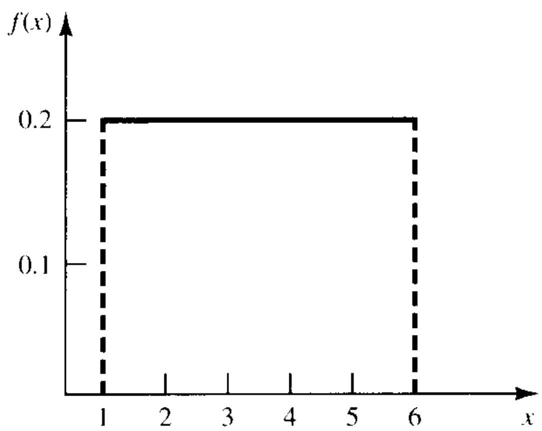
**The uniform distribution:**

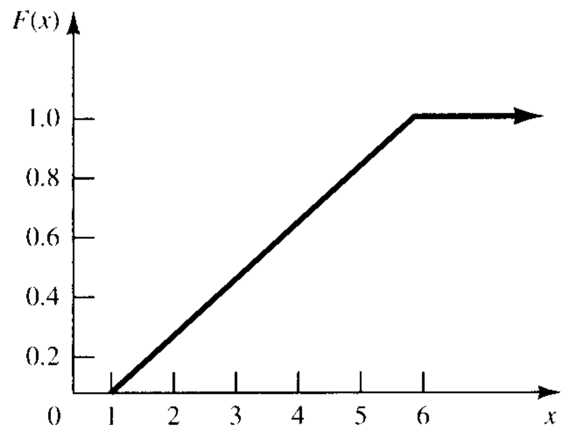
$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

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

$$E[X] = \frac{a+b}{2}$$

$$Var[X] = \frac{(b-a)^2}{12}$$




**The normal distribution:**

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[ -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \right] \quad \text{where} \quad -\infty < x < \infty$$

Letting $y = \dfrac{x-\mu}{\sigma}$ or $x = \sigma y + \mu$ we get a normalized distribution :

$$f_Y(y) = f_X(x)\frac{dx}{dy} \qquad \text{(to be proven)}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \exp\left[ -\frac{1}{2} y^2 \right] \sigma = \frac{1}{\sqrt{2\pi}} \exp\left[ -\frac{y^2}{2} \right]$$

**Mean and variance of Normal distribution:**

$$E[Y] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{1}{2}y^2}\, dy$$

$$= 0 \quad \text{because the integrand is odd.}$$

$$Var[Y] = E[Y^2] - 0$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{1}{2}y^2}\, dy$$

From a table of integrals: $\displaystyle\int_{0}^{\infty} u^2 e^{-au^2}\, du = \frac{1}{4a}\sqrt{\frac{\pi}{a}} \rightarrow \frac{\sqrt{2\pi}}{2} \text{ for } a = \frac{1}{2}$

$$Var[Y] = \frac{1}{\sqrt{2\pi}} \quad \sqrt{2\pi} = 1 \quad \text{because the integrand is even.}$$

Now for $X = \sigma Y + \mu$

$$E[X] = \sigma \quad 0 + \mu = \mu$$

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

$$= \sigma^2 E[Y^2] + 2\mu\sigma \quad E[Y] + \mu^2$$

$$= \sigma^2 \quad 1 + 2\mu\sigma \quad 0 + \mu^2$$

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

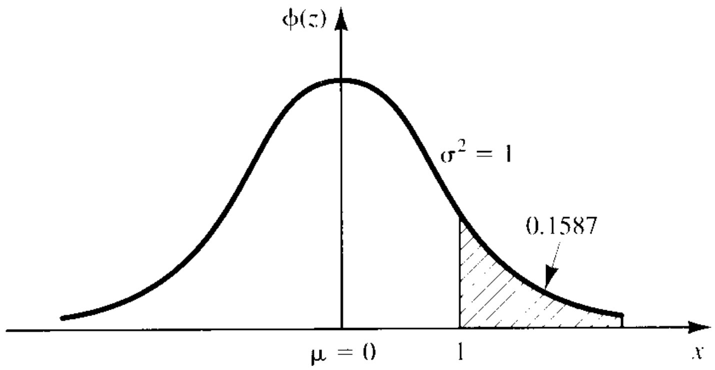
In order to evaluate probabilities of events taken from a Normal distribution we are confronted with the task of calculating probabilities such as:

$$\int_{a}^{b} f_X(x)dx = P(a < X \le b) = P\left(\frac{a-\mu}{\sigma} < Y \le \frac{b-\mu}{\sigma}\right)$$

$$= \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\, dy$$

$$= \int_{-\infty}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\, dy - \int_{-\infty}^{\frac{a-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\, dy$$

This is difficult since there is no closed form solution for the indefinite integral. However, this last form of the normalized integral is tabulated for positive values. (SeeTable A.3 in Banks, Carson, Nelson and Nicol). Since the distribution is symmetrical about the mean, this tabulation suffices for all values of the upper limit.




The shape of the Normal pdf is bell shaped and extends to minus and plus infinity. The figure shows the normalized Normal with mean 0 and standard deviation 1. It is interesting to note that the area beyond the value $x = 1$ is only 0.1587. So that the region within one standard deviation from the mean represents 68.3% of the possible samples from the distribution. The region outside of 3 standard deviations from the mean is generally considered to be negligible.

**Homework assignment:** (from Banks, Carson, Nelson and Nicol.)

1. Do Exercise 6 in Chapter 5. What assumption(s) do you have to make?
2. Do Exercise 9 in Chapter 5.
3. Do Exercise 22 in Chapter 5.
4. Do Exercise 37 in Chapter 5.