Lecture 6

Example: Sum of two independent random variables

$Z=X+Y$

$$\int_a^b f_z(z)dz = P(a < Z \leq b)$$

$$= P(a < X + Y \leq b)$$
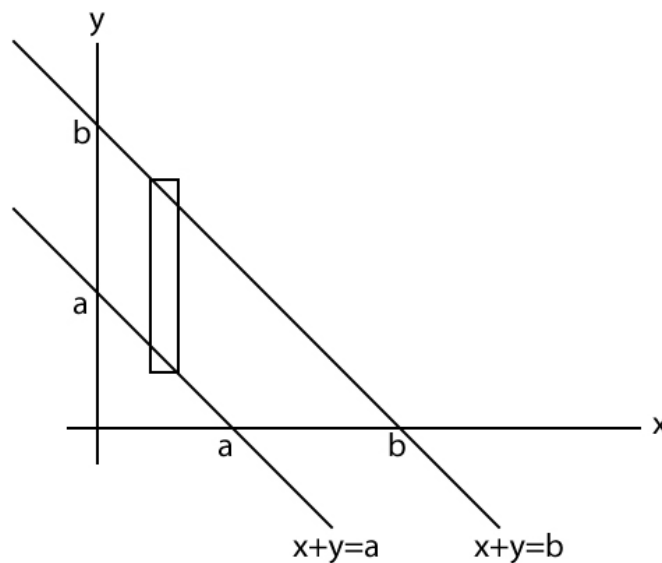
$$= P(a - X < Y \leq b - X)$$

$$= \lim_{dx \to 0} \sum_x P(x < X \leq x + dx)P(a - x < Y \leq b - x)$$

$$= \lim_{dx \to 0} \sum_x f_x(x)dx \int_{a-x}^{b-x} f_y(y)dy$$

$$= \int_{-\infty}^{\infty} f_x(x)dx \int_{a-x}^{b-x} f_y(y)dy$$

We can reach this same point by just integrating the joint probability density function for X and Y over the region for which the event is true.

In the interior strip, the event $a < z \leq b$ is true.

$$P(a < Z \leq b) = \int_{-\infty}^{\infty} dx \int_{a-x}^{b-x} f_{x,y}(x,y)dy$$

$$= \int_{-\infty}^{\infty} dx f_x(x) \int_{a-x}^{b-x} f_y(y)dy$$

Let $z = x + y,\ dz = dy$

$$= \int_{-\infty}^{\infty} dx f_x(x) \int_{a}^{b} f_y(z-x)dz$$

$$= \int_{a}^{b} \left[\int_{-\infty}^{\infty} f_x(x) f_y(z-x)dx \right] dz$$

This is true for all a,b. Therefore:

$$f_z(z) = \int_{-\infty}^{\infty} f_x(x) f_y(z-x)dx$$

$$= \int_{-\infty}^{\infty} f_y(y) f_x(z-y)dy$$

This result can readily be generalized to the sum of more independent random variables.

$$Z = X_1 + X_2 + \dots + X_n$$

$$f_z(z) = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \dots \int_{-\infty}^{\infty} dx_{n-1} f_{x_1}(x_1) f_{x_2}(x_2) \dots f_{x_{n-1}}(x_{n-1}) f_{x_n}(z - x_1 - x_2 - \dots - x_{n-1})$$

Also, if $W = Y - X$, for X, Y independent:

$$f_w(w) = \int_{-\infty}^{\infty} f_x(x) f_y(w+x)dx$$

$$= \int_{-\infty}^{\infty} f_y(y) f_x(y-w)dy$$

Direct determination of the joint probability density of several functions of several random variables

Suppose we have the joint probability density function of several random variables X, Y, Z, and we wish the joint density of several other random variables defined as functions X, Y, Z.

$$U = \tilde{u}(X,Y,Z)$$

$$V = \tilde{v}(X,Y,Z)$$

$$W = \tilde{w}(X,Y,Z)$$

If $f_{x,y,z}(x,y,z)$ is <u>finite everywhere</u>, the density $f_{u,v,w}(u,v,w)$ can be found directly by the following steps:

1. Evaluate the Jacobian of the transformation from X,Y,Z to U,V,W.

$$J(x,y,z) = \begin{vmatrix} \dfrac{\partial \tilde{u}(x,y,z)}{\partial x} & \dfrac{\partial \tilde{u}(x,y,z)}{\partial y} & \dfrac{\partial \tilde{u}(x,y,z)}{\partial z} \\[2mm] \dfrac{\partial \tilde{v}(x,y,z)}{\partial x} & \dfrac{\partial \tilde{v}(x,y,z)}{\partial y} & \dfrac{\partial \tilde{v}(x,y,z)}{\partial z} \\[2mm] \dfrac{\partial \tilde{w}(x,y,z)}{\partial x} & \dfrac{\partial \tilde{w}(x,y,z)}{\partial y} & \dfrac{\partial \tilde{w}(x,y,z)}{\partial z} \end{vmatrix}$$

2. For every value of u,v,w, solve the transformation equations for x,y,z. If there is more than one solution, get all of them.

$$\left. \begin{aligned} \tilde{u}(X,Y,Z) &= u \\ \tilde{v}(X,Y,Z) &= v \\ \tilde{w}(X,Y,Z) &= w \end{aligned} \right\} \rightarrow \begin{cases} x_i(u,v,w) \\ y_i(u,v,w) \\ z_i(u,v,w) \end{cases}$$

3. Then

$$f_{u,v,w}(u,v,w) = \sum_i \frac{f_{x,y,z}(x_i,y_i,z_i)}{\left| J(x_i,y_i,z_k) \right|}$$

    with x_i, y_i, z_i given in terms of u,v,w.

This approach can be applied to the determination of the density function for m variable which are defined to be functions of n variables ($n>m$) by adding some <u>simple auxiliary variables</u> such as x,y,etc. to the list of m so as to total n variables. Then apply this procedure and finally integrate out the unwanted auxiliary variables.

Example: Product U=XY

To illustrate this procedure, suppose we are given $f_{x,y}(x,y)$ and wish to find the probability density function for the product $U = XY$.

First, define a second random variable; for simplicity, choose $V = X$. Then use the given 3 step procedure.

1. Evaluate the Jacobian:

$$J(x,y) = \begin{vmatrix} y & x \\ 1 & 0 \end{vmatrix} = -x$$

2. Solve the transformation equations:

$$\left. \begin{aligned} xy &= u \\ v &= x \end{aligned} \right\} \Rightarrow \begin{cases} x = v \\ y = \dfrac{u}{x} = \dfrac{u}{v} \end{cases}$$

3. Then find:

$$f_{u,v}(u,v) = \frac{1}{|v|} f_{x,y}(v, \frac{u}{v})$$

To get the density for the product only, integrate out with respect to v.

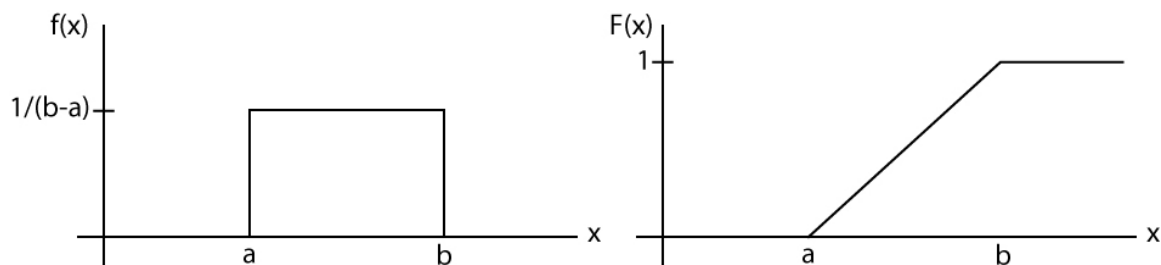$$f_u(u) = \int_{-\infty}^{\infty} \frac{1}{|u|} f_{x,y}(v, \frac{u}{v}) dv$$

If X and Y are independent this becomes

$$f_u(u) = \int_{-\infty}^{\infty} \frac{1}{|v|} f_x(v) f_y\left(\frac{u}{v}\right) dv, \text{ or}$$

$$f_u(u) = \int_{-\infty}^{\infty} \frac{1}{|x|} f_x(x) f_y\left(\frac{u}{x}\right) dx$$

The Uniform Distribution

In our problems we have been using the uniform distribution without having concisely defined it. This is a continuous distribution in which the probability density function is uniform (constant) over some finite interval.

Thus a random variable having a uniform distribution takes values only over some finite interval (a,b) and has uniform probability density over that interval. In what situation does it arise? Examples include part tolerances, quantization error, limit cycles. Often you do not know anything more than that the unknown value lies between known bounds.

$$\bar{X} = \int_a^b x \frac{1}{b-a} dx = \frac{\frac{1}{2}(b^2 - a^2)}{b-a}$$

$$= \frac{1}{2}(a+b)$$

$$\overline{X^2} = \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{3}\frac{1}{b-a}(b^3 - a^3)$$

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

$$\sigma^2 = \overline{X^2} - \bar{X}^2 = \frac{1}{3}(a^2 + ab + b^2) - \frac{1}{4}(a^2 + 2ab + b^2)$$

$$= \frac{1}{12}(a^2 - 2ab + b^2)$$

$$= \frac{1}{12}(b-a)^2$$

$$\sigma = \frac{1}{\sqrt{12}}(b-a)$$

The Binomial Distribution

Outline:
1. Definition of the distribution
2. Determination of the binomial coefficient and binomial distribution
3. Useful relations in dealing with binomial coefficients and factorials
4. The mean, mean square, and variance of the binomial distribution

1. Definition of the distribution

Consider an experiment in which we identify two outcomes: one of which we call success and the other failure. The conduct of this experiment and the observation of the outcome may be called a simple trial. If the trial is then repeated under such circumstances that we consider the outcome on any trial to be independent of the outcomes on all other trials, we have a process frequently called Bernoulli Trials after the man who first studied at length the results of such a process.

The number of successes in n Bernoulli trials is a random discrete variable whose distribution is known as the Binomial Distribution.

Note that the binomial distribution need not refer only to such simple situations as observing the number of heads in n tosses of a coin. An experiment may have a great many simple outcomes – the outcomes may even be continuously

distributed – but for some purposes we may choose to regard some collection of the simple events as success, and employ the binomial distribution to study such successes. For example, in the manufacture of transistors, the resistivity of each crystal is measured after the crystal is drawn. The resistivity of a crystal may be regarded as a random variable taking values over a continuous range, but for the purposes of quality control in the manufacturing process, the occurance of any resistivity value within the specified tolerance may be regarded as success and all other values as failure. From this point of view, the experiment of measuring the resistivity of a crystal has only two outcomes.

How many units must be produced to yield $P=0.95$ (for example) that 100 will be within specifications. Or, for a given yield, how many units must be manufactured to achieve a $P=0.95$ of getting 100 "in-spec" items (not sequentially evaluated).

2. Determination of the binomial coefficient and binomial distribution

The probability of any specified arrangement of k successes and $n\text{-}k$ failures in n independent trials is $p^k q^{n-k}$ where p is the probability of success on any one trial and $q=1\text{-}p$ is the probability of failure. The probability of exactly k successes in n trials is then $p^k q^{n-k}$ times the number of arrangements in which exactly k successes occur. This is due to the fact that the different arrangements are mutually exclusive so their probabilities add.

The probability of <u>exactly</u> k successes in n trials where p is the probability of success on each individual trial is

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} p^k q^{n-k}$$

The number $\binom{n}{k}$ is called the binomial coefficient since it is exactly the coefficient of the kth term in the binomial expansion of $(a+b)^n$.

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

3. Useful relations in dealing with binomial coefficients and factorials

Use the convention that for integral n, $\binom{n}{k} = 0$ for $k<0$ and for $k>n$. Also use the conventions that

$$\binom{n}{0} = 1$$

$$\binom{n}{n} = 1$$

Useful relations involving sums of binomial coefficients:

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}, \quad n \text{ any number}$$

$$\binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{n} = 2^n, \quad n \text{ a positive integer}$$

$$\binom{n}{0} - \binom{n}{1} + \dots \pm \binom{n}{n} = 0, \quad n \text{ a positive integer}$$

4. The mean, mean square, and variance of the binomial distribution

$$
\begin{aligned}
G(s) &= \sum_k p_k s^k \\
&= \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} s^k \\
&= \sum_{k=0}^{n} \binom{n}{k} (ps)^k q^{n-k} \\
&= (ps+q)^n
\end{aligned}
$$

The moments can be derived from this using

$$\frac{dG}{ds} = n(ps+q)^{n-1} p$$

$$\frac{d^2 G}{ds^2} = n(n-1)(ps+q)^{n-2} p^2$$

$$\bar{X} = \frac{dG}{ds}\bigg|_{s=1} = np$$

$$
\begin{aligned}
\overline{X^2} &= \frac{d^2 G}{ds^2}\bigg|_{s=1} + \frac{dG}{ds}\bigg|_{s=1} \\
&= n(n-1)p^2 + np
\end{aligned}
$$

$$
\begin{aligned}
\sigma^2 &= \overline{X^2} - \bar{X}^2 \\
&= n^2 p^2 - np^2 + np - n^2 p^2 \\
&= npq
\end{aligned}
$$

The Poisson Distribution

Outline:
1) Discussion of the significance of the distribution
2) Summarize the derivation of the distribution
3) The Poisson approximation to the Binomial distribution
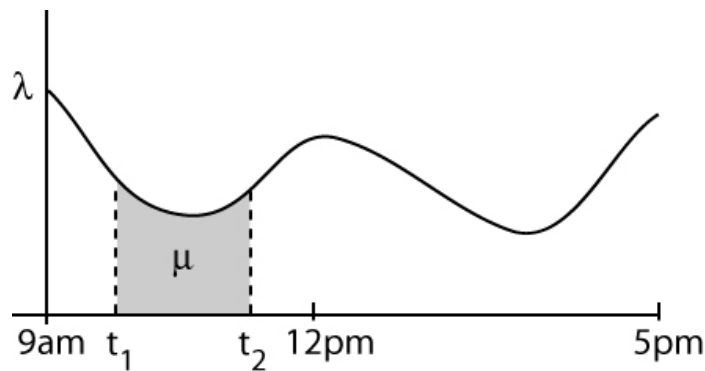
1. Discussion of the significance of the distribution

The Posson distribution is a discrete integer-valued distribution which is of great importance in practical problems. It is employed in the study of <u>telephone traffic</u> through a switchboard, of the <u>emission of B rays</u> from radioactive material, of the <u>emission of electrons</u> from heated filaments and photo-sensitive surfaces, of <u>line surges</u> in a power transmission system due to the throwing of switches, of <u>demands for service</u> in a store or stockroom, and others. Using the statistics of the Poisson distribution, one answers questions related to the number of telephone operators, or stock clerks, or turnstiles in the subway which are required to yield a certain quality of service – where the quality of service is defined in terms of the <u>probability that a customer will have to wait</u>, or the <u>expectation of the time he must wait</u>, or the like. In control system analysis, the Poisson distribution may well lead to an excellent statistical description of the inputs seen by a system, or the occurrence of disturbances, etc.

This distribution applies – in rough terms – to situations in which events occur at random over an interval I and we wish to know the probability $P(k, I, \lambda)$ of the occurrence of exactly k events in the subinterval I_1, of I where λ is the average rate of occurrence of the events – which may possibly vary over the interval. Very often the interval is time: thus we ask the probability that exactly k electrons be emitted in a given length of time.

However, other dimensions may equally well apply – such as intervals of distance. They may even be multidimensional – units of surface area or volume. An example of an <u>interval of volume</u> is given by the statistics of the <u>number of bacteria in volume samples of blood</u>. In the "aerodynamics" of very high altitudes we may very well employ the Poisson distribution to give the probability of occurrence of k atoms in a specified volume.

The conditions under which the Poisson distribution may be expected to apply are given by the assumptions under which the distribution is derived.

Example: Telephone calls throughout a business day

where λ is the average number of phone calls per minute.

$$P(\text{exactly } k \text{ events in an interval } (t_1,t_2)) = P(k) = \frac{1}{k!}\mu^k e^{-\mu}$$

where $\mu = \int_{t_1}^{t_2} \lambda(\tau)d\tau$.

2. Summarize the derivation of the distribution

In an analytic derivation of the distribution, the following assumptions are made:

a. The occurrence of events in all non-overlapping intervals are independent.

b. The probability of the occurrence of 1 event in an infinitesimal interval *dI* is an infinitesimal of the order of *dI*; the probability of occurrence of more than one event in that interval is an infinitesimal of higher order than *dI*.

These assumptions permit the derivation of the distribution; thus if the circumstances of a given situation seem to meet these conditions, the Poisson distribution may be expected to apply.
The derivation proceeds by determining the probabilities for 0 and 1 event which establish an apparent functional form for the distribution, then by proving the form by showing in general that the form holds for 0 and 1 and that if it holds for *n*, it holds also for *n+1*.
The derivation is developed in adequate detail in the text, and leads to the general form of the Poisson distribution:

$$P(k) = \frac{1}{k!}\mu^k e^{-\mu}, \text{ where } \mu = \int_{t_1}^{t_2} \lambda(\tau)d\tau$$

This is the probability of exactly *k* events in the interval (t_1,t_2) in which $\lambda(t)$ is the average rate of occurrence of events over the interval.

If the average rate is the constant λ, and the interval length is written as t, the probability of k events in the interval t becomes

$$P(k) = \frac{1}{k!}(\lambda t)^k e^{-\lambda t}$$

$$\begin{aligned}
G(s) &= \sum_k p_k s^k \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} \mu^k e^{-\mu} s^k \\
&= e^{-\mu} \sum_{k=0}^{\infty} \frac{(\mu s)^k}{k!} \\
&= e^{-\mu} e^{\mu s} \\
&= e^{\mu(s-1)}
\end{aligned}$$