EC-1(Measurement and Instrumentation)Unit 1 Notes (III)

Amarendra Narayan

Distribution functions and their derivation

Probility Distribution

A Probability distribution is a function that assigns a probability value corresponding to each value of a random variable (i.e. For each possible outcome of an experiment.) A probability distribution gives the likelihood of each outcome on the basis of probability theory *before the actual experiment has been done.*

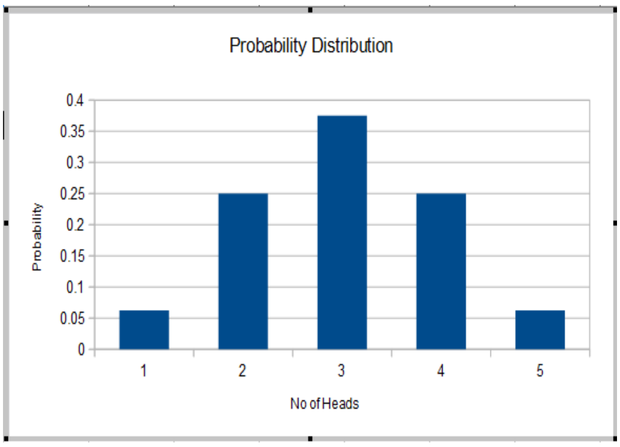
Since probability distribution is a function, it can be represented by a graph with the random variable (independent variable) shown along the x-axis and the probabilities (dependent variable) along the y-axis. We can also express this function using suitable analytical expression where such relation exists. A third way is to depict the function in a table of values.

Example

Suppose that a balanced coin is tossed 4 times. Then there are the following 16 equally likely outcomes: {HHHH, HHHT, HHTH, HTHH, THHH, HHTT, HTHT, THHT, HTTH, THTH, TTHH, HTTT, THTT, TTHT, TTTH, TTTT}. We can make a table for the probabilities of different number of heads like this:

Number of Heads	Frequency of Occurrence	Probability
0	1	$\frac{1}{16}$
1	4	$\frac{4}{16}$
2	6	$\frac{6}{16}$
3	4	$\frac{4}{16}$
4	1	$\frac{1}{16}$

This data can also be shown as a histogram:



Discrete Probability Distribution

A discrete probability distribution is the distribution of probabilities of different values assumed by a discrete random variable. The example given above is of a discrete probability distribution.

Continuous Probability Distribution

A variable is continuous if the values that it assumes have infinitesimal spacings. This means that the difference between any two values of the variable is infinitesimal. If a particular value exists, then we can find another value arbitrarily close to it (as close to it as *we choose*).

A continuous probability distribution is the distribution of probabilities of different values assumed by a continuous random variable. In a continuous distribution the spacing between consecutive values is infinitesimally close. So within even a very small range there is a very large number of values (Theoretically an infinite number of values).

We use probability density function for expressing the probabilities.

Some common probability distributions

When we talk about probability distributions in real life, then depending on the nature of data they can be anything. But when we talk about theoretical probabilities, we make certain simplifying assumptions and then calculate probability based on these assumptions. Depending on the assumptions we can have several distributions. Out of these we will study three which are most useful in our day-to-day situations.

1. Binomial Distribution (Discrete)
2. Gaussian distribution (Continuous)
3. Poisson distribution (Discrete)

Binomial Distribution

There are many situations where a trial is repeated several times (say, n times) and we are interested in finding out the probability of x successes assuming that the probability of success in each trial is p.

For example, the probability of a six coming in a die toss is ⅙. If we toss it 5 times what is the probability of six coming twice?

So, n = 5, x = 2, p = ⅙.

In such cases where the total number of trials is finite and each trial is independent of others, we use Binomial distribition.

$$P(x) = {}^nC_x\, p^x (1-p)^{n-x}$$

Binomial distribution is based on the following assumptions:

1. The number of trials or repetitions of the experiment (n) is fixed.
2. The probability of success is same (p) in each trial.
3. The trials are independent. The outcome of a trial does not affect the probability of success of any other trial.

The Binomial distribution is valid only where the above assumptions hold.

Consider an experiment that has two outcomes. Take one particular outcome, the probability of whose occurrence is *p*. Then the probability of its non-occurrence is $q = 1 - p$. This is the probability of occurrence of the other outcome. Now suppose that we repeat this experiment twice. The trials are independent. There are four possibilities when the order is considered.

Case I :	S S	P = p x p	n = 2	x = 2	$P = p^2$
Case II :	S F	P = p x (1 - p)	n = 2	x = 2	P = p(1-p)
Case III :	F S	P = (1 - p) x p	n = 2	x = 2	P = p(1-p)
Case IV :	F F	P = (1 - p) x (1 - p)	n = 2	x = 2	$P = (1-p)^2$

The probabilities of the different possible outcomes when we ignore order of occurrence is given below:

Occurrences	Probability	Coefficients
2 times	$(p/2)*(p/2)=p^2/4$	1
1 time	$(p*q + q*p)/4 = 2pq/4$	2
None time	$(q/2)*(q/2) = q^2/4$	1

If the trial is repeated 3 times then the outcomes will be as following:

3 times	: $p*p*p = p^3$.	Coefficients	: 1
2 times	: $p*p*q + p*q*p + q*p*p = 3\ p^2q$		: 3
1 time	: $p*q*q + q*q*p + q*p*\ q = 3\ pq^2$		: 3
0 time	: $q*q*q = q^3$		: 1

If the trial is repeated 4 times then the outcomes will be as following:

4 times	: $p*p*p*p = p^4$	Coefficients	: 1
3 times	: $p*p*p*q + p*p*q*p + p*q*p*p + q*p*p*p = 4\ p^3q$		: 4
2 times	: $p*p*q*q + p*q*p*q + p*q*q*p + q*p*q*p + q*q*p*p + q*p*p*q= 6\ p^2q^2$		: 6
1 time	: $p*q*q*q + q*p*q*q + q*q*p*q + q*q*q*p = 4\ pq^3$		: 4
0 time	: $q*q*q*q = q^4$		: 1

We obtain a pattern for coefficients as we increase the number of trials that we can extrapolate. In the triangle given below, each number is the sum of the two numbers immediately above it. This pattern is called Pascal's triangle. It can be further extended if we want to increase the number of trials.

```
                1
               / \
             1     1
            /  \  /  \
          1     2     1
         / \   / \   /  \
       1    3     3     1
      /  \  / \  / \   /  \
    1     4     6    4     1
   /  \  /  \  / \  /  \  /  \
  1    5    10    10    5    1
```

Total Probability = Success probability + Failure probability = 1.

Since the trials are independent, the combined outcome in n trials is the product of the individual outcomes.

Total Probability = (Success probability + Failure probability)*(Success probability + Failure probability)..n times.

If p is the success probability and q is the fdailure probability then

$$1=(p+q)^n={}^nC_0\,p^n q^0+{}^nC_1\,p^{n-1}q^1+{}^nC_2\,p^{n-2}q^2+\ldots+{}^nC_r\,p^{n-r}q^r+\ldots+{}^nC_n\,p^0 q^n$$

The first term consists of the probability for all successes. The second term consists of the probability for one failure and all rest successes, and so on. So the rth term represents the probability for (n-r) successes and r failures.

So we have $1=(p+q)^n=P(n)+P(n-1)+\ldots+P(n-r)+\ldots+P(0)$.

On comparison we see that $P(n-r)={}^nC_r\,p^{n-r}q^r$

Since ${}^nC_r={}^nC_{n-r}$ we have $P(r)={}^nC_r\,p^r q^{n-r}$

Now, $q=1-p$ so that

$$P(n,r)={}^nC_r\,p^r(1-p)^{n-r}$$

The above formula defines the Binomial distribution. Binomial distribution is a discrete distribution. Binomial distribution represents several important processes including distribution of errors in repeated measurements.

If p = q = 0.5, then the distribution is symmetric about n/2. Some of its important parameters are following:

Expectation value $=N\,p$

Variance $\sigma^2=N\,pq$

Moment of skewness $a_3=\dfrac{q-p}{\sqrt{Npq}}$

Moment of Curtosis $a_4=3+\dfrac{1-6\text{pq}}{Npq}$

Mean and Variance of Binomial Random Variables

The probability function for a binomial random variable is

$$P(x)={}^nC_x\,p^x(1-p)^{n-x}$$

This is the probability of having x successes in a series of n independent trials when the probability of success in any one of the trials is p. If X is a random variable with this probability distribution,

$$E(x)=\sum_{x=0}^{n} x\,{}^nC_x\, p^x(1-p)^{n-x}$$
$$=\sum_{x=0}^{n} x\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$$
$$=\sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!}p^x(1-p)^{n-x}$$

since the $x=0$ term vanishes. Let $y=x-1$ and $m=n-1.$ Subtituting $x=y+1$ and $n=m+1$ into the last sum (and using the fact that the limits $x=1$ and $x=n$ correspond to $y=0$ and $y=n-1=m,$ respectively) we get

$$E(x)=\sum_{y=0}^{m}\frac{(m+1)!}{y!(m-y)!}p^{y+1}(1-p)^{m-y}$$
$$=(m+1)p\sum_{y=0}^{m}\frac{m!}{y!(m-y)!}p^{y}(1-p)^{m-y}$$
$$=np\sum_{y=0}^{m}\frac{m!}{y!(m-y)!}p^{y}(1-p)^{m-y}$$

The binomial theorem says that

$$(a+b)^m=\sum_{y=0}^{m}\frac{m!}{y!(m-y)!}a^y b^{m-y}$$

Setting $a=p$ and $b=1-p$ we get

$$\sum_{y=0}^{m}\frac{m!}{y!(m-y)!}p^{y}(1-p)^{m-y}=\sum_{y=0}^{m}\frac{m!}{y!(m-y)!}a^y b^{m-y}=(a+b)^m=(p+1-p)^m=1$$

so that

$$E(X)=np$$

Similarly, but this time using $y=x-2$ and $m=n-2$

$$E(X(X-1))=\sum_{x=0}^{n}x(x-1)\,{}^nC_x\, p^x(1-p)^{n-x}$$
$$=\sum_{x=0}^{n}x(x-1)\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$$
$$=\sum_{x=2}^{\infty}\frac{n!}{(x-2)!(n-x)!}p^x(1-p)^{n-x}$$
$$=n(n-1)p^2\sum_{x=2}^{\infty}\frac{n!}{(x-2)!(n-x)!}p^{x-2}(1-p)^{n-x}$$
$$=n(n-1)p^2\sum_{y=0}^{m}\frac{m!}{(y)!(m-y)!}p^{y}(1-p)^{m-y}$$
$$=n(n-1)p^2(p+(1-p))^m$$
$$=n(n-1)p^2$$

So the variance of X is

$$E(X^2)-E(X)^2=E(X(X-1))+E(X)-E(X^2)=n(n-1)p^2+np-(np)^2$$

Thus

$$\sigma^2=E(X^2)-E(X)^2=np(1-p)$$

Poisson Distribution

Consider a situation in which the probability of occurrence of an event is very small, but there is a possibiity of almost infinite trials; where the probability of occurrence of the next event is independent of earlier occurrences, where the events are so much spaced that the probability of two events in the same slot is negligible. Then we use the Poisson distribution.

$$f(x)=\frac{(np)^x e^{-np}}{x!}$$

If we put $(np)=\lambda$ then we can write $f(x)=\frac{(\lambda)^x e^{-\lambda}}{x!}$

Note : λ can be a fractional number but x is an integer. This is why Poisson distribution is a discrete distribution.

Note :

Some example applications:

1. A Bank receives an average of $\lambda=6$ bad cheques per day. What is the probability that it will receive $x=4$ cheques on a given day?

   Solution: $$f(4)=\frac{6^4 . e^{-6}}{4!}=\frac{(1.296)(0.0025)}{24}=0.135$$

2. Suppose that at a road intersection $\lambda=1.6$ accidents can be expected on any given day. What is the probability that $x=3$ accidents take place on a particular day?

   Solution: $$f(3)=\frac{1.6^3 . e^{-1.6}}{3!}=\frac{(4.096)(0.202)}{6}=0.138$$

3. If the prices of new cars increase on average four times every three years, find the probability of (i) no price hike in a randomly selected 3 year period, (ii) Two price hikes, (iii) Four price hikes, (iv) five price hikes.

4. Given a binomial distribution with n = 28 trials and p = 0.025. Use the Poisson approximation to the binomial to find

   (i) $P(r\geq 3)$ (ii) $P(r<5)$ (iii) $P(r=9)$.

5. Given $\lambda=6.1$ for a Poisson distribution find (i) $P(x\leq 3)$ (ii) $P(x\geq 2)$ (iii) $P(x=6)$ (iv) $P(1\leq x\leq 4)$

6. On the average five birds hit Patna TV Tower every week and get killed. What is the probability that more than three birds get killed in a week?

Proof of Poisson distribution: Let us suppose that events are occurring at different instances of time. Let δt denote a particular small duration of time in which an event can either occur or not occur. We have assumed that the probability of two events occuring in the same time slot δt is negligible (zero).

Let $P_1(\delta t)=\lambda\delta t$ be the probability of occurrence of one (1) event in the time duration between t and $t+\delta t$.

$$P_1(\delta t)=\lambda\delta t \tag{1}$$

The probability of non-occurrence of an event during δt is

$$P_0(\delta t)=1-\lambda\delta t \tag{2}$$

If the event has not occurred by the time t, then the probability that it has still not occurred by the time $t+\delta t$ is

$$P_0(t+\delta t)=P_0(t)(1-\lambda\delta t) \quad (3)$$

Or

$$\frac{P_0(t+\delta t)-P_0(t)}{\delta t}=-\lambda P_0(t)$$

If we make δt infinitely small, and call it dt then we can write

$$\frac{dP_0(t)}{dt}=-\lambda P_0(t) \quad (4)$$

On integration we get

$$P_0(t)=e^{-\lambda t}+C \quad (5)$$

If we choose a duration of time zero, then $P_0(0)=1$ so that $C=0$.

Thus

$$P_0(t)=e^{-\lambda t} \quad (6)$$

Now consider the case where $n\geq 1$.

We will look at the probability of occurrence of n events in time $t+\delta t$. To calculate this, we consider the two ways in which this can occur: (1) The event has already occurred by time t, and then it does not occur in the next δt . And (2) It has not occurred up to time t, and occurs during the next δt . The probability of occurrence of n events in time $t+\delta t$ will be the sum of these two probabilities. Thus

$$P_n(t+\delta t)=P_n(t)(1-\lambda\delta t)+P_{n-1}(t)\lambda\delta t \quad (7)$$

In the limit that $\delta t\rightarrow 0$ we have

$$\frac{dP_n(t)}{dt}+\lambda P_n(t)=\lambda P_{n-1}(t) \quad (8)$$

In order to solve this differential equation we need an integrating factor, i.e. a function which, when multiplied with LHS makes it a perfect differential. i.e. We want a function $\mu(t)$ such that

$$\mu(t)\left[\frac{dP_n(t)}{dt}+\lambda P_n(t)\right]=\frac{d}{dt}[\mu(t)P_n(t)] \quad (9)$$

One such function that satisfies this criterion is

$$\mu(t)=e^{\lambda t} \quad (10)$$

This is because

$$\frac{d}{dt}\left[e^{\lambda t}P_n(t)\right]=\lambda e^{\lambda t}P_n(t)+e^{\lambda t}\frac{dP_n(t)}{dt} \quad (11)$$

Using this in (8) we obtain

$$\frac{d}{dt}\left[e^{\lambda t}P_n(t)\right]=\lambda e^{\lambda t}P_{n-1}(t) \quad (12)$$

When we put n = 1 we get

$$\frac{d}{dt}\left[e^{\lambda t}P_1(t)\right]=\lambda e^{\lambda t}P_0(t) \tag{13}$$

Since by (6) we have $P_0(t)=e^{-\lambda t}$ we can write

$$\frac{d}{dt}\left[e^{\lambda t}P_1(t)\right]=\lambda e^{\lambda t}e^{-\lambda t}=\lambda \tag{14}$$

Now we integrate this.

$$e^{\lambda t}P_1(t)=\int \lambda\, dt=\lambda t+C \tag{15}$$

If we choose $t=0,$ the probability to find one event in zero time duration is zero. So that $P_0(0)=0.$ So $C=0.$ (16)

Thus

$$P_1(t)=\lambda t e^{-\lambda t} \tag{17}$$

We can also write this as $P_1(t)=\frac{\lambda t}{1!}e^{-\lambda t}$.

Let us now apply the induction method to generalize this result to arbitrary n . Let us assume that

$$P_n(t)=\frac{(\lambda t)^n}{n!}e^{-\lambda t} \tag{18}$$

is true for a particular n . In particular, we know that (as per 17) it is true for $n=1.$ Using (12) we can write

$$\frac{d}{dt}\left[e^{\lambda t}P_{n+1}(t)\right]=\lambda e^{\lambda t}P_n(t)=e^{\lambda t}\lambda\frac{(\lambda t)^n}{n!}e^{-\lambda t}=\frac{\lambda(\lambda t)^n}{n!} \tag{19}$$

When we integrate this over t we get the following:

$$e^{\lambda t}P_{n+1}(t)=\int\frac{\lambda(\lambda t)^n}{n!}dt=\frac{(\lambda t)^{n+1}}{(n+1)!}+C \tag{20}$$

Imposing the boundary condition $P_{n+1}(0)=0$ gives us $C=0.$ So we finally get it that

$$P_{n+1}(t)=\frac{(\lambda t)^{n+1}}{n+1!}e^{-\lambda t} \tag{21}$$

We now have two assertions:

1. From Equation (17) we see that Equation (18) is true for $n=1.$
2. If Equation (18) is true for $n,$ it is also true for $n=1.$ This is shown by equation (21).
3. Thus Equation (18) is true for n = 1, 2, 3, ... and consequently for all higher values of n.

Equation (18) gives us the expression for Poisson distribution. To recall, Poisson distribution is applicable for situations where the following conditions are satisfied:

1. The probability of occurrence of an event is very small
2. There is a possibiity of almost infinite trials
3. Where the probability of occurrence of the next event is independent of earlier occurrences

4. Where the events are so much spaced that the probability of two events in the same slot is negligible.

Poisson distribution as a special case of Binomial distribution:

The Binomial distribution can be written as following:

$$P(N,n)=\frac{N!}{(N-n)!\,n!}\,p^n(1-p)^{N-n}$$

We will assume that $p\rightarrow 0$ and $N\rightarrow\infty$.

We will make two approximations:

1. $(1-p)^{N-n}\approx e^{-np}$

   This can be shown in the following way:

   $$\ln\left[(1-p)^{N-n}\right]=(N-n)\ln(1-p)$$

   If $p\ll 1$ then using the series expansion $\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\ldots\ for -1<x<1$ we write $\ln(1-p)\approx -p$ so that the RHS becomes $(N-n)\ln(1-p)\approx(N-n)(-p)\approx -Np$

   So

   $$(1-p)^{N-n}\approx e^{-Np} \qquad (1)$$

2. Using Stirling Approximation we write

   $$\ln\left[\frac{N!}{(N-n)!}\right]=N\ln N-N-(N-n)\ln(N-n)+(N-n)$$

   For $n\ll 1$ $$\ln(N-n)=\ln N+\ln\left(1-\frac{n}{N}\right)=\ln N-\frac{n}{N}$$

   or $$(N-n)\left(\ln N-\frac{n}{N}\right)=N\ln N-n\ln N-n+\frac{n^2}{N}$$

   If $n\ll N$ then $n^2/N\rightarrow 0$ so $\ln(N!/(N-n)!)\approx n\ln N$

   Thus we have $$\frac{N!}{(N-n)!}\approx N^n$$

   On putting these together we get

   $$P(N,n)=\frac{(Np)^n e^{-Np}}{n!}$$

   If we put Np as λ then we get

   $$P(N,n)=\frac{\lambda^n e^{-\lambda}}{n!}$$

   This is the expression for Poisson distribution.

Sum of Probabilities:

The sum of probabilities for all n is obtained as:

$$\sum_{n=0}^{\infty} P_n(\nu)=\sum_{n=0}^{\infty} \frac{\nu^n}{n!} e^{-\nu}=e^{-\nu}\sum_{n=0}^{\infty} \frac{\nu^n}{n!}=e^{-\nu} e^{\nu}=1$$

Mean and Standard Deviation:

The mean is given by

$$\langle n\rangle=\sum_{n=0}^{\infty} nP(n)=e^{-\mu}\mu\sum_{n=1}^{\infty} \frac{\mu^{n-1}}{(n-1)!}=\mu e^{-\mu} e^{\mu}=\mu$$

The variance is obtained as following:

$$\langle n^2\rangle=\sum_{n=0}^{\infty} n^2 P(n)=\mu^2+\mu$$

which gives

$$\sigma^2=\langle (n-\mu)^2\rangle=\langle n^2\rangle-\mu^2=\mu$$

Gaussian Distribution

So far we have studied the Binomial distribution and the Poisson distribution. These were discrete distributions. This means that the probability values are available at discrete intervals of the corresponding random variable. Thus in the Binomial distribution the different values of the random variable r for a given value of n are discretely spaced. In the Poisson distribution the different values of the random variable x for a given value of λ are discretely spaced (x = 0,1,2,3,4,...). But in the Gaussian distribution the random variable x is not discrete. It is continuous. This means that two values of x can be as closely spaced as you desire. Let us see what difference it makes.

Probabilities of a continuous random variable: If x is a continuous random variable, then the probability of having different values for it is distributed continuously over some range of x, then (i) the probability for any particular value will be negligible. (ii) The probability of the value of x falling within a particular range will be finite and will be in general proportional to the interval for small intervals. Thus

$$P(x_1< x< x_2)\alpha(x_2-x_1)$$

Let $$P(x_1< x< x_2)=f(x).(x_2-x_1)$$

Then the function $f(x)$ is called Probability density function.

If the interval (x_2-x_1) is very small then we can call it as dx and $P(x_1< x< x_2)$ as dP. Then

$$dP=f(x)dx$$

$$P(x_1< x< x_2)=\int_{x_1}^{x_2} dP=\int_{x_1}^{x_2} f(x)dx$$

We sometimes call $P(x_1< x< x_2)$ as Probability mass function. Thus Probability density is probability per unit interval of the random variable x over a small interval.

A derivation from basic principles:

Consider throwing a dart at the origin of the Cartesian plane. You are *aiming* at the origin, but random errors in your throw will produce varying results. We assume that:

- the errors do not depend on the orientation of the coordinate system.

- errors in perpendicular directions are independent. This means that being too high doesn't alter the probability of being off to the right.
- large errors are less likely than small errors.

In Figure 1, below, we can argue that, according to these assumptions, your throw is more likely to land in region A than either B or C, since region A is closer to the origin. Similarly, region B is more likely that region C. Further, you are more likely to land in region F than either D or E, since F has the larger area and the distances from the origin ;are approximately the same.

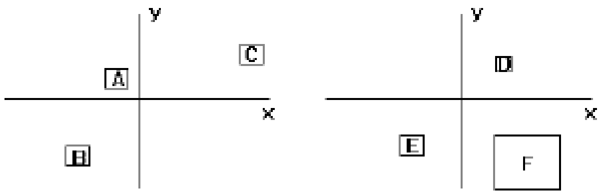


Figure 1

Determining the Shape of the Distribution

Consider the probability of the dart falling in the vertical strip from x to $x+\Delta x$. Let this probability be denoted $p(x)\Delta x$. Similarly, let the probability of the dart landing in the horizontal strip from y to $y+\Delta y$ be $p(y)\Delta y$. We are interested in the characteristics of the function p. From our assumptions, we know that function p is not constant. In fact, the function p is the normal probability density function.

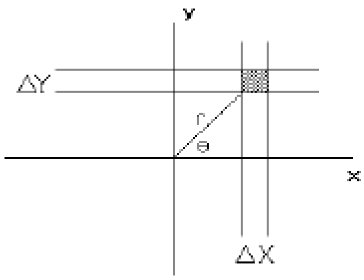


Figure 2

From the independence assumption, the probability of falling in the shaded region is $p(x)\Delta x \times p(y)\Delta y.$ Since we assumed that the orientation doesn't matter, that any region r units from the origin with area D*x* ×D*y* has the same probability, we can say that

$$p(x)\Delta x \times p(y)\Delta y = g(r)\Delta x \Delta y.$$

This means that

$$g(r) = p(x)p(y).$$

Differentiating both sides of this equation with respect to θ, we have

$$0 = p(x)\frac{dp(y)}{d\theta} + p(y)\frac{dp(x)}{d\theta}$$

as g is independent of orientation, and therefore, θ.

Using $x = r\cos(\theta)$ and $y = r\sin(\theta)$, we can rewrite the derivatives above as

$$0 = p(x)p'(y)(r\cos\theta) + p(y)p'(x)(-r\sin\theta)$$

Rewriting again, we have

$$0 = p(x)p'(y)x - p(y)p'(x)y.$$

This differential equation can be solved by separating variables,

$$\frac{p'(x)}{x\,p(x)} = \frac{p'(y)}{y\,p(y)}$$

This differential equation is true for any x and y, and x and y are independent. That can only happen if the ratio defined by the differential equation is a constant, that is, if

$$\frac{p'(x)}{x\,p(x)}=\frac{p'(y)}{y\,p(y)}=C$$

On solving the equation $\frac{p'(x)}{x\,p(x)}=C$ we get $\ln p(x)=\frac{Cx^2}{2}+c$ or $p(x)=A\,e^{\frac{Cx^2}{2}}$

We now use the assumption that the probability of large errors is less than the probability for small errors. This gives us

$$p(x)=A\,e^{\frac{kx^2}{2}}$$

where k is positive. This argument has given us the basic form of the Gaussian distribution. The curve is bell-shaped with maximum value at $x=0$ and points of inflection at $x=\frac{\pm 1}{\sqrt{k}}$. Now we need to find out appropriate values of A and k.

Since p is a probability distribution the total area under the curve must be 1. Thus we need to evaluater the following integral equation:

$$\int_{-\infty}^{\infty} A\,e^{\frac{-kx^2}{2}}\,dx=1$$

The integrand is symmetric about $x=0$. so the integrals over positive and negative parts will be equal. We can consider just the positive part of the integral and write

$$\int_{0}^{\infty} e^{\frac{-kx^2}{2}}\,dx=\frac{1}{2\mathrm{A}}$$

Since x is a dummy variable, the same will be true with x replaced by y . Thus,

$$\left(\int_{0}^{\infty} e^{\frac{-kx^2}{2}}\,dx\right)\left(\int_{0}^{\infty} e^{\frac{-ky^2}{2}}\,dy\right)=\frac{1}{4\mathrm{A}^2}$$

Since x and y are independent, we can rewrite the product as a double integral.

$$\iint_{0}^{\infty} e^{\frac{-k}{2}(x^2+y^2)}\,dy\,dx=\frac{1}{4\mathrm{A}^2}$$

This double integral is now converted into polar co-ordinates. We write

$$\iint_{0}^{\infty} e^{\frac{-k}{2}(x^2+y^2)}\,dy\,dx=\int_{0}^{\frac{\pi}{2}}\int_{0}^{\infty} e^{\frac{-k}{2}r^2}\,r\,dr\,d\theta$$

Let us now make the following substitution: Let $u=\frac{kr^2}{2}$ Then $du=kr\,dr$ and we have

$$\int_{0}^{\frac{\pi}{2}}\int_{0}^{\infty} e^{\frac{-k}{2}r^2}\,r\,dr\,d\theta=\int_{0}^{\frac{\pi}{2}}\frac{-1}{k}\left[\int_{0}^{\infty} e^{u}\,du\right]d\theta=\int_{0}^{\frac{\pi}{2}}\frac{d\theta}{k}=\frac{\pi}{2\mathrm{k}}$$

Now, since $\frac{1}{4\mathrm{A}^2}=\frac{\pi}{2\mathrm{k}}$ so $A=\sqrt{\frac{k}{2\pi}}$. The probability distribution is

$$p(x)=\sqrt{\frac{k}{2\pi}}\;e^{\frac{-kx^2}{2}}$$

Now we would like to express the value of $p(x)$ in terms of mean and variance.

The mean μ is defined as $\mu=\int_{-\infty}^{\infty} x\,p(x)\,dx.$ Since the integrand is odd, the mean is zero.

The variance σ^2 is the value of the integral $\int_{-\infty}^{\infty}(x-\mu)^2\,p(x)\,dx$. This is an even function. So we can just integrate from 0 to ∞ and double the value obtained.

$$\sigma^2=2\sqrt{\frac{k}{2\pi}}\int_0^{\infty} x^2 e^{\frac{-k}{2}x^2}\,dx$$

We can evaluate this integral by parts with $u=x$ and $dv=x\,e^{-k\frac{x^2}{2}}$ to generate the expression

$$2\sqrt{\frac{k}{2\pi}}\left(\left[\lim_{m\to\infty}\frac{-x}{k}e^{\frac{-k}{2}x^2}\right]_0^M+\frac{1}{k}\int_0^{\infty} e^{\frac{-k}{2}x^2}\,dx\right)$$

Now, $\lim_{m\to\infty}\frac{-x}{k}e^{\frac{-k}{2}x^2}=0$ and $\frac{1}{k}\int_0^{\infty} e^{\frac{-k}{2}x^2}\,dx=\frac{1}{k}\frac{\sqrt{2\pi}}{2\sqrt{k}}$

So $2\sqrt{\frac{k}{2\pi}}\frac{1}{k}\int_0^{\infty} e^{\frac{-k}{2}x^2}\,dx=2\frac{\sqrt{k}}{\sqrt{2\pi}}.\frac{1}{k}.\frac{\sqrt{2\pi}}{2\sqrt{k}}=\frac{1}{k.}$

So $k=\frac{1}{\sigma^2}$

The normal probability density function

Now we can obtain the normal probability density function from the three basic assumptions.

$$p(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2}$$

This general expression for the normal distribution with mean μ and standard deviation σ is obtained by shifting this distribution horizontally.

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