

CHANGE OF PROBABILITY MEASURE

Part1: Normal Variables

Ali Al-Hayki*

✉ ali.hayki@gmail.com

🌀 [alihazki](#)

in [ali-al-hayki](#)

01/02/2020

*Use footnote for providing further information about author (webpage, alternative address)—*not* for acknowledging funding agencies.

ABSTRACT

The purpose of this paper is to help students and professionals, who have introductory knowledge of quantitative finance, to develop more advanced knowledge and skills that one needs in order to follow advanced literature, and to build/appreciate models used in trading, pricing, and risk management. The transition from introductory to advanced is the most rewarding, but is one of the hardest steps in the development of a quant and I am aiming to help you with this transition. This paper will particularly develop the concept of change of probability measure using normal distribution as an example. It introduces the technical concepts under the simplest possible settings, which will then be applied to finite stochastic process and Brownian motion.

Keywords First keyword · Second keyword · More

Contents

1	Introduction	3
2	Normal Distribution	3
2.1	Q Measure	5
2.2	Probability Measure Verification	6
2.3	Relationship Between P and Q Distributions	7
2.4	Continuity	7
2.5	Measure Theory	8
A	Moment Generating Function of the Normal Distribution	12
A.1	The Normal PDF is Valid	12

1 Introduction

In this paper we are going to explain the change of of probability measure which is one of the simplest yet one of the most powerful concepts with far-reaching consequences. It is also one of the most commonly tested concepts in job interviews. We are going to develop the concept using standard normal distribution. Later we will change the settings to a finite dimensional stochastic process which will be easy because by then we would have learned the tricks of the trade. This will lead us nicely to the Cameron-Martin-Girsanov Theorem for Brownian Motion. We will then give some examples of its uses such as changing the drift of a process, evaluating truncated conditional expectation using derivations of Black-Scholes formula as an example and importance sampling.

2 Normal Distribution

Lets start with a simple Normally distributed random variable. Assume that a variable under a given probability measure, P , follows a standard normal distribution:

$$X \text{ under } P \sim N[0, 1]$$

Recall the formula of the *density function* a standard normal is:

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

the distribution function which gives the probability of the random variable taking values less than a certain value, say X , can be written as the integral of the density function:

$$P(x) = P[X \leq x] = \int_{-\infty}^x p(t)dt \quad (2)$$

we can write the write the relationship in differential form as well:

$$dP(x) = p(x)dx \quad (3)$$

Now the distribution function is essentially a probability measure because it gives probability of events which in the present case are intervals for example we can write the probability that the value of the random variable is between alpha and beta as follows:

$$P[\alpha \leq X \leq \beta] = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{1}{2}x^2} dx \quad (4)$$

So the change of the distribution function on the same random variable is nothing but an example of a probability measure change, to see this let's assume that we shift the interval by a constant:

$$P[\alpha - \mu \leq X \leq \beta - \mu] = \frac{1}{\sqrt{2\pi}} \int_{\alpha - \mu}^{\beta - \mu} e^{-\frac{1}{2}x^2} dx \quad (5)$$

Essentially we shifted the lower and the upper integration limits by the same constants leaving the random variable X untouched - **Leaving the variable unchanged is the main trick!** Now lets preform a change of variable so the interval of integration gets aligned to equation 4. *Think of this as an attempt to determine how much shift in the probability of the original interval has been induced by the given shift of the integration interval.* Lets define a new variable $y = x + \mu$. when x equals the lower integration limit $= \alpha - \mu$ then Y becomes α :

$$y = \alpha - \mu + \mu = \alpha \quad (6)$$

Analogously when x equals the upper integration limit $= \beta - \mu$ then Y becomes β :

$$y = \beta - \mu + \mu = \beta \quad (7)$$

Finally, $x = \mu$ and taking the derivative with respect to y we get $dx = dy$. Making the substitutions into equation 5.

$$P[\alpha \leq X \leq \beta] = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{1}{2}(y-\mu)^2} dy \quad (8)$$

So the **intervals are the same** now but the **probability has shifted**. We can call the event $P[\alpha \leq X \leq \beta]$ as event A . Then the shifted event $P[\alpha - \mu \leq X \leq \beta - \mu]$ can be written as the probability of $A - \mu$.

We recognise this as the probability of a normal variable with mean μ .

Now lets analyse this simple relationship in more detail:

$$P[A - \mu] = P[\alpha \leq X \leq \beta] = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{1}{2}(y-\mu)^2} dy \quad (9)$$

As y is just a dummy variable, we can switch back to x to make the comparison clearer

$$= \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{1}{2}(x-\mu)^2} dy \quad (10)$$

Expand the squares:

$$= \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{1}{2}(x^2 - 2x\mu + \mu^2)} dx \quad (11)$$

Spiting the exponential into 2 terms

$$= \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{1}{2}x^2} e^{x\mu - \frac{1}{2}\mu^2} dx \quad (12)$$

The aim is isolate the standard normal density which is obvious now and we can write it in terms of the standard normal density which we denoted with small P :

$$= \int_{\alpha}^{\beta} e^{x\mu - \frac{1}{2}\mu^2} p(x) dx \quad (13)$$

Recalling the relationship between the distribution function in differential form:

$$= \int_{\alpha}^{\beta} e^{x\mu - \frac{1}{2}\mu^2} dP(x) \quad (14)$$

We can now represent equation 14 in a few alternative forms which will help us with the terminology.

Form 1: The interval can be denoted in terms of its name A (i.e. $\int_{\alpha}^{\beta} = \int_A$)

$$= \int_A e^{x\mu + -\frac{1}{2}\mu^2} dP(x) \quad (15)$$

Form 2: The interval can also be 'indicated' using an indicator function which take a value of 1 for values of the variable within the indicated interval and 0 otherwise $1_A : X \rightarrow \{0, 1\}$ (i.e. the indicator only activates if a certain condition is true).²

As the indicator function nullifies the domain that is outside the interval of interest, we can write the limit of integration to go from minus infinity to plus infinity

$$= \int_{-\infty}^{\infty} e^{x\mu + -\frac{1}{2}\mu^2} 1_{[\alpha, \beta]} dP(x) \quad (16)$$

now this is nothing but the expected value of the exponential function times the indicator function of the event.

$$= E^P[e^{x\mu + -\frac{1}{2}\mu^2} 1_A] \quad (17)$$

The exponential is a random variable, as it is a function of x , we will denote it with the symbol Z ,

$$= E^P[Z(x) 1_A] \quad (18)$$

It was all very simple and trivial but will help us when we get into more complicated settings. Try to remember the steps no matter how trivial they look in the current settings.

We previously said that this shift of the interval can be viewed as leading to a change of probability so lets use it to define a new probability measure,

we can use any of the alternative forms discussed. The 2 forms that will be mostly relevant are equations 20 and 17.

$$Q[A] = \int_A e^{x\mu + -\frac{1}{2}\mu^2} dP(x) \quad (19)$$

2.1 Q Measure

$$Q[A] = \int_A e^{x\mu + -\frac{1}{2}\mu^2} dP(x) = E^P[Z(x) 1_A] \quad (20)$$

Lets calculate the cumulative probability of a variable taking a value below a ,

$$Q[X \leq a] = E^P[e^{x\mu + -\frac{1}{2}\mu^2} 1_{X \leq a}] \quad (21)$$

and writing the expectation in terms of the integration with respect to the probability,

²If an indicator function is over a subset of random variables, then the expectation of the indicator function is the probability of that subset occurring.

$$Q[X \leq a] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{x\mu - \frac{1}{2}\mu^2} 1_{X \leq a} e^{-\frac{1}{2}x^2} dx \quad (22)$$

reflecting the indicator function in the integration limit,

$$Q[X \leq a] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{x\mu - \frac{1}{2}\mu^2} e^{-\frac{1}{2}x^2} dx \quad (23)$$

Combine the 2 exponentials,

$$Q[X \leq a] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{1}{2}(-2x\mu + \mu^2 + x^2)} dx \quad (24)$$

Which we recognise as a complete square,

$$Q[X \leq a] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{1}{2}(x-\mu)^2} dx \quad (25)$$

Now this is just the distribution function of a normal variable with mean μ so it is a valid distribution function. A distribution function is a probability measure as can be easily verified by recalling the definition of the probability measure.

2.2 Probability Measure Verification

Probability Measure: λ

- $0 \leq \lambda(E_i) \leq 1$: The probability measure of an event has to be between 0 and 1 which the distribution function satisfies.
- $\lambda(\phi) = 0, \lambda(\Omega) = 1$: The probability of the null set and the whole sample must be 0 and 1 respectively which the distribution function satisfies.
- $\{E_i\} = \{E_1, E_2, \dots\}$: If we have a list of disjoint events then the probability of their union is the sum of their probabilities

$$\lambda\left(\bigcup E_i\right) = \sum \lambda(E_i) \quad (26)$$

Our Q definitely satisfies this, so the Q is definitely a probability measure.

So we started with a probability distribution P :

$$X \text{ under } P \sim N[0, 1] \quad (27)$$

and we showed that shifting the interval induces a new distribution Q :

$$X \text{ under } Q \sim N[\mu, 1] \quad (28)$$

2.3 Relationship Between P and Q Distributions

Let's analyse the relationship between the 2 distributions now.

$$Q[A] = \int_A e^{x\mu - \frac{1}{2}\mu^2} dP(x) \quad (29)$$

We can write the Q probability of the X taking values between α and y as follows:

$$Q[\alpha \leq X \leq y] = \int_{\alpha}^y e^{x\mu - \frac{1}{2}\mu^2} dP(x) \quad (30)$$

We can write the left-hand side as the difference between the values of the distribution function.

$$Q(y) - Q(\alpha) = \int_{\alpha}^y z(x) dP(x) \quad (31)$$

Where we use the symbol Z to represent the exponential function. Now this is very similar to the fundamental theorem of calculus which establishes a connection between the two central operations of calculus - integration and differentiation. We produce it in a convenient form here:

$$F(y) - F(a) = \int_a^y f(x) dx \quad (32)$$

The small f function f is the derivative of the capital F function F

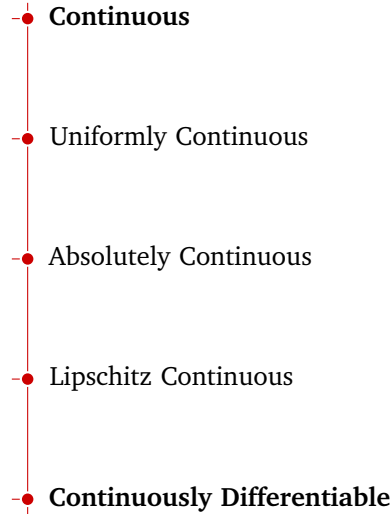
$$\frac{dF(y)}{dy} = f(y) \quad (33)$$

Hence the capital F function F is the anti derivative or the primitive or the integral of the small f . You can easily see the relationship between the probability measures Q and P through the Z function. The question we face now are what conditions are needed on the function and its derivative, that's the capital F function and its derivative the small f so that both sides of the above equation are defined and equal. If the derivative is continuous then the relationship between the derivative and the function is given by the fundamental theorem of calculus holds, but we would like the conditions in terms of the function itself. Requiring that the function be just continuous in the sense of the usual definition is not sufficient. It indeed sounds strange that we cannot recover the function by integrating its derivative and the derivative by differentiating its integral even if the function is continuous so the trouble must be in the continuity definition and we tighten the continuity definition then things might work out .

2.4 Continuity

Recall that only a small subset of the continuous functions are differentiable between the two there are many shades of continuity:

Continuity Spectrum



All these definitions call a function continuous if a small change in the input leads to a small change in the output but they assess them smallness differently. The requirement get tighter as one moves from the top to bottom and as such the number of functions which satisfies the criteria reduces as one moves from the top to the bottom. If the derivative is continuous then we are at the bottom end of the continuum and the relationship between the derivative and the function that is given by the fundamental theorem holds. If the function is just continuous then we are at the top end of the continuum and this requirement is not sufficient for the relationship to hold. So the right answer must be in the middle. If the function is absolutely continuous then the fundamental theorem holds. Absolute continuity imposes stricter requirements from the basic continuity to get a feel for the differences in the requirements we quickly outline the definition of absolute continuity and relate it to the usual epsilon-delta definition. We will not need this technical definition as things simplify when one moves from the real line to the general measure spaces but it is good to know for context. The definition goes in the usual epsilon-delta style. A function $F(y)$ is absolutely continuous on $[a, b]$, if for every $\epsilon > 0$, there is a δ such that for every finite collection of disjoint intervals $\{(a_i, b_i), (a_i, b_i), \dots, (a_i, b_i)\}$ if the absolute variation of the inputs is less than Delta then the absolute variation of the function outputs is less than epsilon:

$$\sum_{i=1}^n |b_i - a_i| < \delta, \quad \text{then} \quad \sum_{i=1}^n |F(b_i) - F(a_i)| < \epsilon \quad (34)$$

Setting n equal to 1 gives a definition of uniform continuity and if one is allowed to choose Delta separately based on X then one gets the usual epsilon delta definition of local continuity. The definition simplifies when one switches to measures.

2.5 Measure Theory

A measure Q is said to be absolutely continuous with respect to another measure P ($Q \ll P$) if whenever measure P assigns 0 probability to an event then Q also assigns zero probability to the same event (i.e. if $P(E) = 0$ then

$Q(E) = 0$). In simple terms the null set under P is also the null set under Q . Notice the absolute continuity is represented by the double less than symbol \ll . An easy way to remember this is to say that Q is dominated by P in that the null set under P must be a null set under Q . As we know the probability measure is indeed a measure so this definition applies to probability measures as well. For the probability measure defined by the continuous distribution functions on the real line that we have been discussing so far. The 2 definitions that is the complicated looking definition above and a simple definition are equivalent, but this simple looking definition is the one that applies in more general settings. Just for completeness the reason this definition does not contain the usual epsilon-delta is because a general measure is not necessary finite and if the measure happens to be finite which should be the case for the probability measure then one recast the definition in terms of the familiar epsilon delta (i.e. if $P(E) < \delta$, then $Q(E) < \epsilon$). Though this definition is not commonly used, it is easy to check that the probability measure Q is absolutely continuous with respect to the measure P and given the similarity of the connection between the two probability measure to the fundamental theorem. It wont come as a big surprise if we call the function Z the derivative of the probability measure Q with respect to the probability measure P :

$$\frac{dQ(y)}{dP(y)} = z(y) \quad (35)$$

The derivative is called the Radon–Nikodym derivative in the technical literature and we shall be calling $z(y)$ the Radon–Nikodym derivative from now onwards. the so called Radon–Nikodym theorem goes a step further. If a measure Q is absolutely continuous with another measure P then it guarantees the existence and uniqueness of the function z or the derivative of the measure with respect to the other. If $Q \ll P$, then there is z such that:

$$Q[A] = \int z(x) dP(x) \quad \frac{dQ(x)}{dP(x)} = z(x) \quad (36)$$

It is also called density for the obvious reasons. If P is absolutely continuous with respect to Q then we say that the two measures are equivalent and one can then switch between the two effortlessly as the Radon Radon–Nikodym applies to measures and our measure being a probability measure has to satisfy some extra conditions such as the total probability must be 1. We need to verify that the function z us indeed a probability density function so we will have to check that it is non-negative and that its expected value is 1 which is the same thing as checking that the total probability sums is up to 1.

$$z(x) = e^{x-\mu-\frac{1}{2}\mu^2} \geq 0 \quad (37)$$

as function 38 is exponential it is non-negative by default. Now for the second requirement, let's calculate its expected value under the measure p :

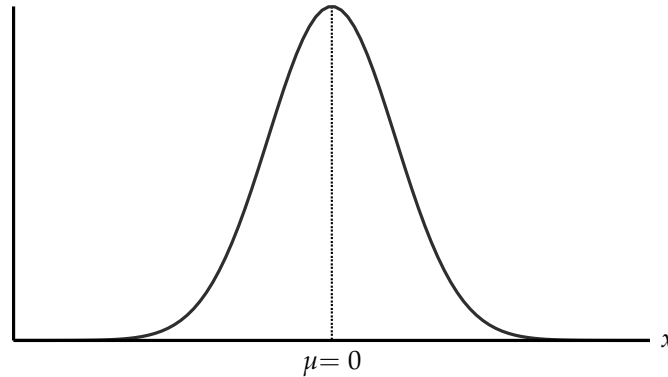
$$E^P \left[e^{x-\mu-\frac{1}{2}\mu^2} \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{x-\mu-\frac{1}{2}\mu^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-\mu)^2} dx = 1 \quad (38)$$

Thus the total probability under Q will be 1 as we can verify by substituting the Radon Radon–Nikodym derivative:

$$E^P [Z(x)] = \int_{-\infty}^{\infty} \frac{dQ(x)}{dP(x)} dP(x) = \int_{-\infty}^{\infty} dQ(x) \quad (39)$$

This also explains the relationship between the expected value under P and Q :

Figure 1: Standard Normal Distribution



$$E^P [Z(x)h(x)] = \int_{-\infty}^{\infty} \frac{dQ(x)}{dP(x)} h(x) dP(x) = \int_{-\infty}^{\infty} h(x) dQ(x) = E^Q [h(x)] \quad (40)$$

That is the expected value under P of Z times a function is equal to the expected value under Q of the function.

Let's see it visually. We have a standard normal variable $X_{\text{under } P} \sim N[0, 1]$ whose density looks like figure 2.5. The familiar bell-shaped function. Now let's create an exact copy of the x-axis and let's define our new probability measure through the Radon–Nikodym derivative:

$$X \text{ under } Q \sim N[1.5, 1] \quad (41)$$

$$Q[A] = \int_A z(x) dP(x) \quad (42)$$

$$z(x) = e^{1.5x - \frac{1}{2}(1.5)^2} \quad (43)$$

In figure 2.5, you can see that the same variable now has a different distribution under the measure Q . The distribution has shifted to the right and the variable now has a mean of 1.5. Function 43, shows where the mean comes from. Going over the X you can see the same interval now has a different probability.

The green rectangle in figure 2.5 shows the probability of a small interval under the original measure P . The blue rectangle in figure 2.5 shows the probability of the same interval under the measure Q (to verify this take a ruler and align it on the edges of the interval). You can see that the probability has shifted to the right. The question for you is how would you equate the expectation under the two measures? Also how can you identify the standard normal under the measure Q ? Would it be X minus something?

Figure 2: Standard Normal Distribution - shifted $\mu + 1.5$

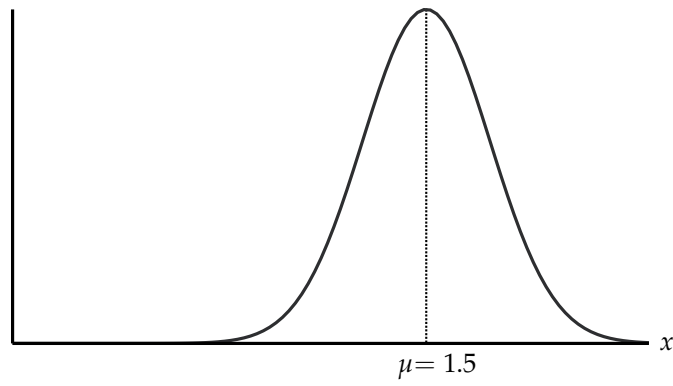


Figure 3: Standard Normal Distribution Interval

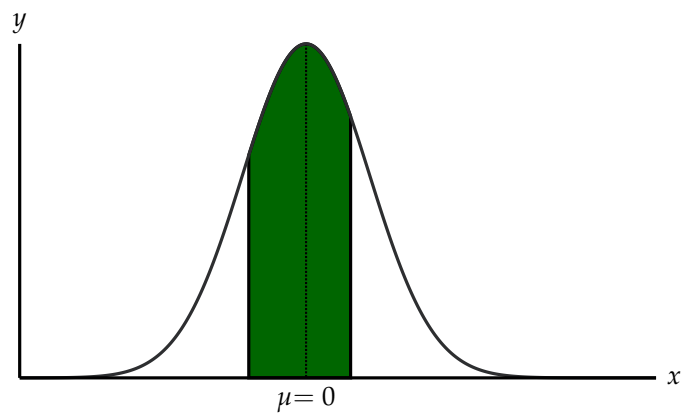
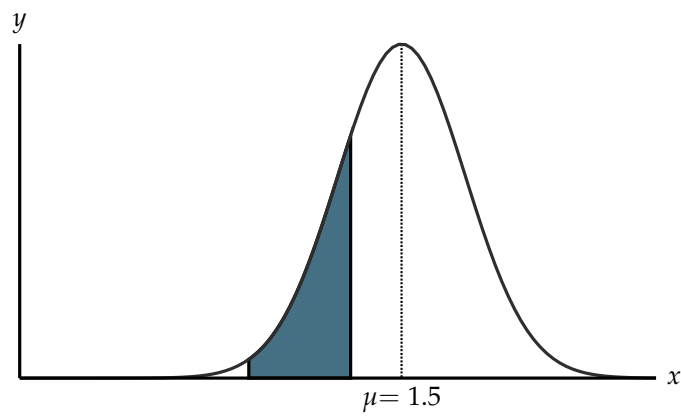


Figure 4: Standard Normal Distribution Interval



A Moment Generating Function of the Normal Distribution

Recall that the probability density function of a normally distributed random variable x with mean of $E(x) = \mu$ and a variance of $V(x) = \sigma^2$ is:

$$N(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (44)$$

Our objective is to find the moment generating function which corresponds to equation 44. To begin, let us consider the case where $\mu = 0$ and $\sigma^2 = 1$. Then we have a standard normal, denoted by $N(z; 0, 1)$, and the corresponding moment generating function is defined by,

$$M_z(t) = E(e^{zt}) = \int e^{zt} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \quad (45)$$

$$A = kjpk \quad (46)$$

A.1 The Normal PDF is Valid

Recall that the probability density function of a normal random variable is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (47)$$

for $-\infty < x < \infty$, $-\infty < \mu < \infty$, and $0 < \sigma < \infty$. Also recall that in order to show that the normal pdf is a valid pdf, we need to show that $f(x)$ is always positive and if we integrate over the entire support, we get 1. **Showing that $f(x)$ is positive**

The standard deviation σ is defined to be positive. The square root of 2π is positive and the natural exponential function is positive. When you multiply positive numbers together, you get a positive number.

Showing that $f(x)$ integrates to 1

Let's define I to be the integral that we are trying to find. That is:

$$I = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx \quad (48)$$

Our goal is to show that $I = 1$. Now, if we change variables with:

$$w = \frac{x - \mu}{\sigma} \quad (49)$$

Our integral I becomes:

$$I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(w)^2} dw \quad (50)$$

Now, squaring both sides, we get:

$$I^2 = \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x)^2} dx \right) \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y)^2} dy \right) \quad (51)$$

REFERENCES

And, pulling the integrals together, we get:

$$I^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} e^{-\frac{1}{2}y^2} dx dy \quad (52)$$

Now, combining the exponents we get:

$$I^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2+y^2)} dx dy \quad (53)$$

Converting to polar coordinates with:

$$x = r \cos \theta \text{ and } y = r \sin \theta \quad (54)$$

we get:

$$I^2 = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \left(\int_0^{\infty} e^{-\frac{r^2}{2}} r dr \right) d\theta \quad (55)$$

Now, if we do yet another change of variables with:

$$u = \frac{r^2}{2} \text{ and } du = r dr \quad (56)$$

our integral I becomes:

$$I^2 = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \left(\int_0^{\infty} e^{-u} du \right) d\theta \quad (57)$$

Evaluating the inside integral, we get:

$$I^2 = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \left(- \lim_{b \rightarrow \infty} [e^{-u}]_{u=0}^{u=b} \right) d\theta \quad (58)$$

and, finally, completing the integration, we get:

$$I^2 = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} -(0 - 1) d\theta = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} d\theta = \frac{1}{\sqrt{2\pi}} (2\pi) = 1 \quad (59)$$

Okay, so we've shown that $I^2 = 1$. Therefore, that means that $I = +1$ or $I = -1$. But, we know that I must be positive, since $f(x) > 0$. Therefore, I must equal 1. Our proof is complete.

References

- [1] Ali Al Hayki. Real-time segmentation of on-line handwritten arabic script. In *Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on*, pages 417–422. IEEE, 2014.