Derivation of the univariate normal distribution

This derivation is a slight adaption from the one given by Dan Teague. Here we elaborate all the steps and incorporate the notion of the expected value.

Before we can start the derivation we should first state some mild assumptions. Consider that we are aiming at the origin of the Cartesian plane with darts. We assume the following:

1. The deviation of the darts does not depend on the orientation of the coordinate system.
2. Deviations in orthogonal directions are independent. This means if we deviate a lot in one direction the probabilities of the other direction are not influenced.
3. Large deviations are less likely than small deviations.

Consider that we are throwing an dart, the probability that it falls in the interval $[x, x + \Delta x]$, under the assumption that Δx is very small (almost infinitesimal), is given by:

$$\int_{x}^{x+\Delta x} p(x)dx \approx p(x)\Delta x$$

For the perpendicular direction this is exactly the same, giving $p(y)\Delta y$. Figure 1 shows this situation, where the shaded area is particular bounded region of which we want to calculate the probability. Under the independence assumption of perpendicular direction we can state that the probability of the bounded region is $p(x)p(y)\Delta x\Delta y$.

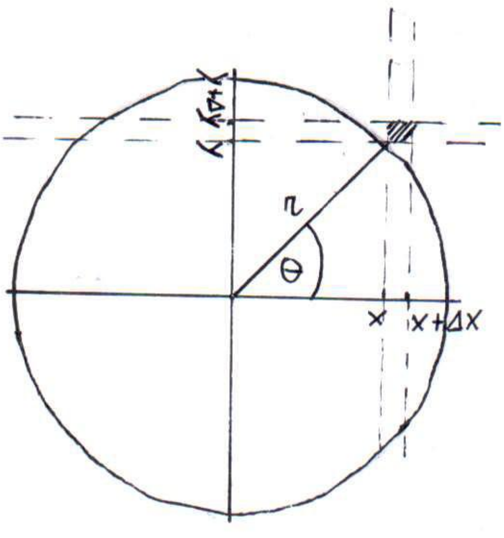

Figure 1: Cartesian plane

It can of course happen that we are not aiming at the origin but instead aim at an offset of the origin identified by the constant μ. This way the probability of our translated bound region becomes:

$$p(x + \mu)p(y + \mu)\Delta x\Delta y$$

We have assumed that the orientation doesn't matter, and that any similar region r units away from the origin with the area $\Delta x\Delta y$ has the same probability. We can state:

$$p(x + \mu)p(y + \mu)\Delta x\Delta y = g(r)\Delta x\Delta y$$

This implies that

$$g(r) = p(x+\mu)p(y+\mu)$$

Where the function $g(r)$ is not dependent on the angle, as every probability is the same. Figure 1 also illustrates a circle with radius r where all the values of the function are the same. Hence this function is only dependent on the radius.

For the values x and y we can also substitute their respective polar conversions, x=$rcos(\theta)$ and $y = rsin(\theta)$. Differentiation with respect to the angle θ on both sides using the product rule and chain rules gives:

$$0 = p(x+\mu)p'(y+\mu)rcos(\theta) - p(y+\mu)p'(x+\mu)rsin(\theta)$$

Converting back to the ordinary Cartesian domain gives:

$$p(y+\mu)p'(x+\mu)y = p(x+\mu)p'(y+\mu)x$$

This differential equation can be solved by separating variables:

$$\frac{p'(x+\mu)}{xp(x+\mu)} = \frac{p'(y+\mu)}{yp(y+\mu)} \quad \forall x, y \in \mathbb{R}$$

We can know that we can pick x and y independently and the differential equation still holds true. This can only be iff:

$$\frac{p'(x+\mu)}{xp(x+\mu)} = \frac{p'(y+\mu)}{yp(y+\mu)} = C$$

Solving:

$$\frac{p'(x+\mu)}{xp(x+\mu)} = C$$
$$\frac{p'(x+\mu)}{p(x+\mu)} = Cx$$
$$\ln p(x+\mu) = \frac{Cx^2}{2} + c$$
$$p(x+\mu) = Ae^{C\frac{x^2}{2}}$$

We assumed that larger deviations are less likely we can thus claim that C must be negative such that the probability density decreases when the deviation from μ becomes larger. We would now like to find the probability density function which is only influenced by x. We do this by:

$$p(x) = p\big((x-\mu)+\mu\big) = Ae^{-C\frac{(x-\mu)^2}{2}}$$

One of the fundamental properties is that the integral over the probability distribution must be equal to one. This structure gives a way to solve for the constant A.

$$A\int_{-\infty}^{\infty} e^{-C\frac{(x-\mu)^2}{2}} dx = 1$$

$$\int_{-\infty}^{\infty} e^{-C\frac{(x-\mu)^2}{2}} dx = \frac{1}{A}$$

We now use the substitution rule:

$$\alpha = x - \mu$$
$$\frac{d\alpha}{dx} = 1$$
$$\int_{-\infty}^{\infty} e^{-C\frac{\alpha^2}{2}} d\alpha = \frac{1}{A}$$

We can clearly see that the integrand is symmetric, hence we can do the following:

$$\int_{0}^{\infty} e^{-C\frac{\alpha^2}{2}} d\alpha = \frac{1}{2A}$$

This integral is hard to evaluate but we can do the following:

$$\int_{0}^{\infty} e^{-C\frac{\alpha^2}{2}} d\alpha \int_{0}^{\infty} e^{-C\frac{\beta^2}{2}} d\beta = \frac{1}{4A^2}$$

Due to the independence of α and β, and the Fubini theorem, we can state:

$$\int_{0}^{\infty}\int_{0}^{\infty} e^{-C\frac{(\alpha^2+\beta^2)}{2}} d\alpha d\beta = \frac{1}{4A^2}$$

To solve the left hand side of the equation we must change the variables of the double integral. We do this by the following theorem.

Let D and D^* be elementary regions in (respectively) the xy-plane and the uv-plane. Suppose $T: \mathbb{R}^2 \to \mathbb{R}^2$ is a coordinate transformation of class C^1 that maps D on D^* in a one-on-one fashion. If $f: D \to \mathbb{R}$ is an integrable function and we use the transformation T to make the substitution $x = x(u, v)$ and $y = y(u, v)$, then:

$$\int\int_D f(x,y) dxdy = \int\int_{D^*} f(x(u,v), y(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| dudv$$

The Jacobian of the transformation T, conveniently denoted by:

$$\frac{\partial(x,y)}{\partial(u,v)}$$

Is the determinant of the derivative matrix $DT(u, v)$. That is for the 2D case,

$$\frac{\partial(x,y)}{\partial(u,v)} = \det D\,T(u,v) = det\begin{bmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{bmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}$$

Using polar coordinates we can make the following transformation:

$$(x, y) = T(r, \cos\theta) = (r\cos\theta, r\sin\theta)$$

The Jacobian for this change of variables is:

$$\frac{\partial(x,y)}{\partial(r,\theta)} = det\begin{bmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{bmatrix} = r(\cos\theta)^2 + r(\sin\theta)^2 = r$$

This eventually results in the following integral:

$$\int_0^{\frac{\pi}{2}}\int_0^{\infty} r\,e^{-C\frac{r^2}{2}}dr\,d\theta = \frac{1}{4A^2}$$

Using the substitution rule:

$$\varphi = \frac{Cr^2}{2}$$
$$\frac{d\varphi}{dr} = Cr$$
$$\frac{d\varphi}{Cr} = dr$$
$$\frac{1}{C}\int_0^{\frac{\pi}{2}}\int_0^{\infty} e^{-\varphi}d\varphi d\theta = \frac{1}{4A^2}$$
$$\frac{1}{C}\int_0^{\frac{\pi}{2}}(-\lim_{t\to\infty} e^{-t} + e^0)d\theta = \frac{1}{4A^2}$$
$$\frac{1}{C}\int_0^{\frac{\pi}{2}} d\theta = \frac{1}{4A^2}$$
$$\frac{\pi}{2C} = \frac{1}{4A^2}$$
$$A = \pm\sqrt{\frac{C}{2\pi}}$$

For the solution of A we only are interested in the real valued ones. Thus, we omit the complex solutions as we want the probability density function to become a real function. We now must evaluate if the solution of A must be negative or positive. The answer is quite intuitive. If we would take the negative solution for A the following could happen:

$$-\sqrt{\frac{C}{2\pi}}e^{-C\frac{(x-\mu)^2}{2}}\Bigg|_{x=\mu} = -\sqrt{\frac{C}{2\pi}}$$

One of the fundamental properties of probability density distributions is that its function is non-negative:

$$p(x) \geq 0 \;\; \forall x \in \mathbb{R}$$

Hence we can omit the negative solution of A, giving the result:

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

Now we should calculate the expected value of this distribution to find the constant C.

$$E[X] = \sqrt{\frac{C}{2\pi}} \int_{-\infty}^{\infty} x e^{-C\frac{(x-\mu)^2}{2}} dx$$

Using the substitution rule:

$$\alpha = x - \mu$$
$$\frac{d\alpha}{dx} = 1$$
$$= \sqrt{\frac{C}{2\pi}} \int_{-\infty}^{\infty} (\alpha + \mu) e^{-C\frac{\alpha^2}{2}} d\alpha = \sqrt{\frac{C}{2\pi}} \left(\int_{-\infty}^{\infty} \alpha e^{-C\frac{\alpha^2}{2}} d\alpha + \mu \int_{-\infty}^{\infty} e^{-C\frac{\alpha^2}{2}} d\alpha \right)$$

The evaluation of the first integral term in the last equation is very easy to evaluate. We see that the integrand is a multiplication of an odd and an even function, resulting in an odd function. A symmetric integral over an odd function is equal to zero. The second integral we have solved above, we now have the result:

$$= \sqrt{\frac{C}{2\pi}} \left(0 + 2\mu \sqrt{\frac{\pi}{2C}} \right) = \sqrt{\frac{C}{2\pi}} \sqrt{\frac{2\pi}{C}} \mu = \mu$$

We know want to calculate the variance of this distribution:

$$\sigma^2 = E[(X - E[X])^2) = E[X^2] - (E[X])^2 = E[X^2] - \mu^2$$
$$\sigma^2 + \mu^2 = \sqrt{\frac{C}{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-C\frac{(x-\mu)^2}{2}} dx$$

Using the substitution rule:

$$\alpha = x - \mu$$
$$\frac{d\alpha}{dx} = 1$$
$$= \sqrt{\frac{C}{2\pi}} \int_{-\infty}^{\infty} (\alpha + \mu)^2 e^{-C\frac{\alpha^2}{2}} d\alpha$$

$$= \sqrt{\frac{C}{2\pi}}\left(\int_{-\infty}^{\infty} \alpha^2 e^{-C\frac{\alpha^2}{2}} d\alpha + 2\mu \int_{-\infty}^{\infty} \alpha e^{-C\frac{\alpha^2}{2}} d\alpha + \mu^2 \int_{-\infty}^{\infty} e^{-C\frac{\alpha^2}{2}} d\alpha \right)$$

These integrals are yet again easy to evaluate using intuitive properties. The middle integral has an integrand which is an odd function, resulting that this integral is equal to zero. The third integral we have seen before, these properties result in:

$$\sigma^2 + \mu^2 = \sqrt{\frac{C}{2\pi}} \int_{-\infty}^{\infty} \alpha^2 e^{-C\frac{\alpha^2}{2}} d\alpha + \mu^2 \sqrt{\frac{C}{2\pi}} \sqrt{\frac{2\pi}{C}}$$
$$\sigma^2 = \sqrt{\frac{C}{2\pi}} \int_{-\infty}^{\infty} \alpha^2 e^{-C\frac{\alpha^2}{2}} d\alpha$$

Note that the integrand of this integral is symmetric, so we can write:

$$\sigma^2 = 2\sqrt{\frac{C}{2\pi}} \int_{0}^{\infty} \alpha^2 e^{-C\frac{\alpha^2}{2}} d\alpha$$

We can solve this integral by making use of integration by parts:

$$f(\alpha) = \alpha$$
$$f'(\alpha) = 1$$
$$g'(\alpha) = \alpha e^{\frac{-C\alpha^2}{2}}$$
$$g(\alpha) = -\frac{1}{C} e^{\frac{-C\alpha^2}{2}}$$
$$= 2\sqrt{\frac{C}{2\pi}} \left(\lim_{t\to\infty} -\frac{t}{C} e^{\frac{-Ct^2}{2}} + \frac{1}{C} \int_{0}^{\infty} e^{\frac{-C\alpha^2}{2}} d\alpha \right)$$

The integral we have evaluated multiple times now and can be substituted with the answer:

$$= 2\sqrt{\frac{C}{2\pi}} \lim_{t\to\infty} -\frac{t}{C} e^{\frac{-Ct^2}{2}} + 2\sqrt{\frac{C}{2\pi}} \frac{1}{C} \frac{\sqrt{2\pi}}{2\sqrt{C}}$$
$$= 2\sqrt{\frac{C}{2\pi}} \lim_{t\to\infty} -\frac{t}{C} e^{\frac{-Ct^2}{2}} + \frac{1}{C}$$

We only need to solve the limit and then we have the answer:

$$-\frac{1}{C} \lim_{t\to\infty} t e^{\frac{-Ct^2}{2}} = -\frac{1}{C} \lim_{t\to\infty} \frac{t}{e^{\frac{Ct^2}{2}}}$$

We should recognize this limit as an indeterminate form $\frac{\infty}{\infty}$, hence we can use l'Hopital rule:

$$= -\frac{1}{C^2} \lim_{t\to\infty} \frac{1}{t e^{\frac{Ct^2}{2}}}$$

To solve this limit, we shall make use of the squeeze theorem. We state that:

$$0 \leq \frac{1}{te^{\frac{Ct^2}{2}}} \leq \frac{1}{e^t} \quad t \geq 1$$

We know that:

$$\lim_{t\to\infty} \frac{1}{e^t} = 0$$

Hence :

$$\lim_{t\to\infty} \frac{1}{te^{\frac{Ct^2}{2}}} = 0$$

Resulting in:

$$\sigma^2 = \frac{1}{C}$$
$$C = \frac{1}{\sigma^2}$$

We now can give you the probability distribution of the univariate normal distribution characterized by its mean and variance, μ and σ^2 respectively.

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

So you see, starting with the assumptions that the large deviations are less likely than small deviations and the symmetry assumption we derived the univariate normal distribution with its expected value at μ (the translated origin of the Cartesian plane in the beginning).