

KEY WORDS: PROBABILITY DISTRIBUTION FUNCTION, GAUSSIAN DISTRIBUTION

KEY CONCEPTS: KNOW HOW TO SHOW THAT $f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ IS A PDF.

20.1 Probability and Expectation

Recall a function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ defines a **probability distribution function** on \mathbb{R}^n if $f(x) \geq 0$ for all x and

$$\int \int \cdots \int_{\mathbb{R}^n} f dV = 1.$$

Underlying the probability distribution function is a probability measure defined by some experiment which gives an outcome to each point in \mathbb{R}^n . If X is the outcome of the experiment, then the probability distribution function gives

$$\mathbb{P}(X \geq \lambda) = \int \int \cdots \int_E f dV$$

where $E = \{x \in \mathbb{R}^n : X \geq \lambda\}$ – so E is the set of points in \mathbb{R}^n which witness the outcome of the experiment being at least λ . A basic example of a probability distribution function is the **uniform distribution**. Here we define $f(x, y) = 1$ for $(x, y) \in D$ and $D = [0, 1] \times [0, 1]$ and $f(x, y) = 0$ otherwise. Then $f(x, y) \geq 0$ for all $(x, y) \in \mathbb{R}^2$ and $\iint_D f dA = 1$, so f is indeed a probability distribution function. Now suppose the experiment we perform is to measure the distance from a random point in D to the origin. Let X be the outcome of this experiment, so that $X(x, y) = \sqrt{x^2 + y^2}$. Now consider the event $X \geq 1$. The points which witness this outcome are the points $(x, y) \in D$ such that $x^2 + y^2 \geq 1$. Then we see that

$$\mathbb{P}(X \geq 1) = \iint_E f dA = \int_0^1 \int_{\sqrt{1-x^2}}^1 1 dy dx = 1 - \frac{\pi}{4}.$$

So the chance that a random point ends up at distance at least one from the origin in the unit square is $1 - \frac{\pi}{4}$.

Next we can ask, on average, where will a random point be? Given a probability distribution function f , and a function $X = g(x)$ measuring the outcome of an experiment, the **expected value** of X is the improper integral $\mu(X) = \iint \cdots \int g f dV$ when it exists and the standard deviation, denoted $\sigma(X)$, is defined by $\sigma(X)^2 = \iint \cdots \int g^2 f dV - \mu(X)^2$. For example, suppose we want the expected distance of the randomly dropped point from the origin. Since the distance is $g(x, y) = \sqrt{x^2 + y^2}$, we want the expected value of $X = g(x, y)$. This is

$$\mu(X) = \iint_D \sqrt{x^2 + y^2} f dA = \int_0^1 \int_0^{\sqrt{1-x^2}} \sqrt{x^2 + y^2} dy dx = \frac{1}{3}(\sqrt{2} + \operatorname{arcsinh} 1) \approx 0.765.$$

The standard deviation is found from

$$\sigma^2(X) = \iint_D (x^2 + y^2) f dA - \mu(X)^2 = \frac{\sqrt{2}}{3} - \frac{1}{9}(\sqrt{2} + \operatorname{arcsinh}1)^2.$$

20.2 Gaussian distribution

One of the fundamental probability distributions in statistics is the [Gaussian](#) or [standard normal](#) distribution. The probability distribution function is

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

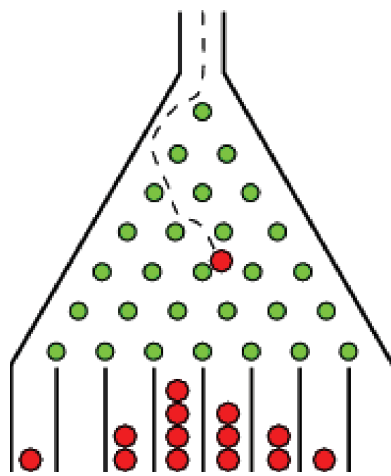
for $x \in \mathbb{R}$. A remarkable fact is that if we perform a sequence of n independent experiments, each having average outcome μ and standard deviation equal to σ , and X is the sum of the values of the experiments, then as $n \rightarrow \infty$,

$$Z = \frac{X - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} N(0; 1).$$

This means that the probability distribution function of Z is asymptotically the Gaussian distribution, regardless of the distributions of the individual experiments. For example, suppose we pick a random number from $\{-1, 1\}$ a total of n times and independently on different picks. We stipulate that 1 and -1 are picked with probability $\frac{1}{2}$ – for example by tossing a fair coin. Then the expected value of each experiment is clearly $\mu = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$ – this is the average outcome. So in the above setup, we have $\mu = 0$. The standard deviation is $\sigma = \frac{1}{2}(1)^2 + \frac{1}{2}(-1)^2 - \mu^2 = 1$. Now if X is the sum of the outcomes, then what we stated above says that X/\sqrt{n} has, asymptotically, the Gaussian distribution. This is known in probability theory as the [Central Limit Theorem](#). That means that if λ is any real number, and Z is defined as above, then

$$\lim_{n \rightarrow \infty} \mathbb{P}(Z \geq \lambda) = \frac{1}{\sqrt{2\pi}} \int_{\lambda}^{\infty} e^{-x^2/2} dx.$$

A nice way to visualize this is via a [Galton Board](#).



A Galton Board
2

Imagine beads dropping from a point directly above the top pin in the Galton Board. At each level, a bead dropping down on to a pin is equally likely to fall to the left or to the right of the pin. The beads eventually fall into bins placed between the bottom row of pins. Now for example, the chance that a bead falls into the leftmost bin is 2^{-n} if there are n rows of pins, because it has to fall to the left at every level, and each of these events has a probability of $\frac{1}{2}$ of occurring, independently over n levels. The middle bins have a higher probability. The central limit theorem predicts that the shape of the pile of beads that we get as the number of beads and number of levels tends to infinity is defined by the curve $y = e^{-x^2/2}$ with appropriate normalization and the point $x = 0$ in the middle.

We have yet to prove that $f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ is a probability distribution function. Clearly $f(x) \geq 0$. It remains to show

$$\int_{\mathbb{R}} f dx = 1.$$

First we compute

$$\iint_{\mathbb{R}^2} e^{-(x^2+y^2)/2} dA.$$

Transforming to polar co-ordinates, this is

$$\int_0^{2\pi} \int_0^{\infty} e^{-r^2/2} r dr d\theta$$

using the change of variables theorem. Making the substitution $w = r^2/2$, we get

$$\int_0^{2\pi} \int_0^{\infty} e^{-w} dw = 2\pi.$$

Now we use a general fact that if $g(x)$ and $h(y)$ are integrable functions on \mathbb{R} , then

$$\iint_{\mathbb{R}^2} g(x)h(y)dA = \left(\int_{\mathbb{R}} g(x)dx \right) \cdot \left(\int_{\mathbb{R}} h(y)dy \right).$$

In our case, we will apply this with $g(x) = e^{-x^2/2}$ and $h(y) = e^{-y^2/2}$ to get

$$2\pi = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)/2} dy dx = \int_{\mathbb{R}} e^{-x^2/2} dx \int_{\mathbb{R}} e^{-y^2/2} dy = \left(\int_{\mathbb{R}} e^{-z^2/2} dz \right)^2.$$

Therefore

$$\int_{\mathbb{R}} e^{-z^2/2} dz = \sqrt{2\pi}$$

and this shows that $\int_{\mathbb{R}} f(x) = 1$, and f is a probability distribution function.