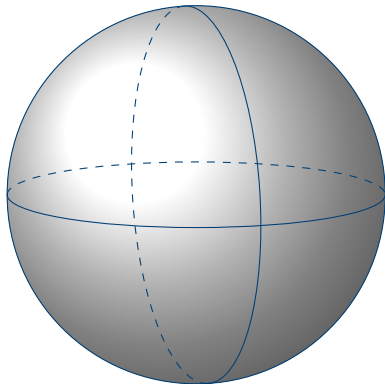


The Borel–Kolmogorov paradox

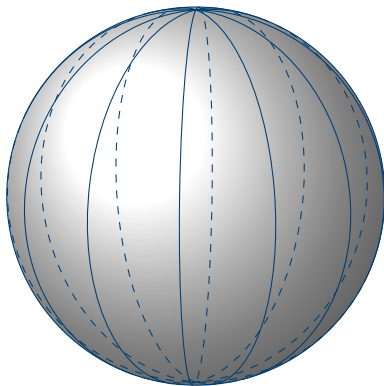
Yarin Gal

yg279@cam.ac.uk

Suppose that a random variable has a uniform distribution on a unit sphere.



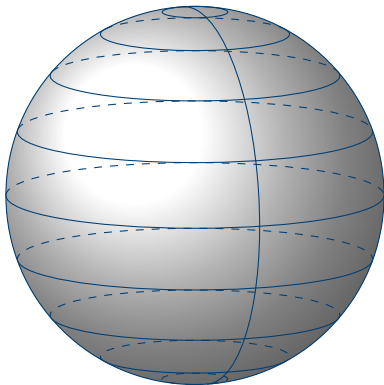
Suppose that a random variable has a uniform distribution on a unit sphere.



Choosing a point uniformly:

1. Choose longitude λ uniformly from $[-\pi, \pi]$

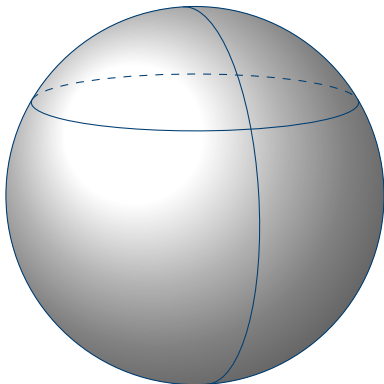
Suppose that a random variable has a uniform distribution on a unit sphere.



Choosing a point uniformly:

1. Choose longitude λ uniformly from $[-\pi, \pi]$
2. Choose latitude ϕ from $[-\frac{\pi}{2}, \frac{\pi}{2}]$ with density $\frac{1}{2} \cos \phi$.

Suppose that a random variable has a uniform distribution on a unit sphere.



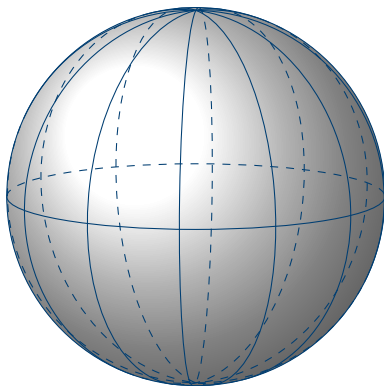
Choosing a point uniformly:

1. Choose longitude λ uniformly from $[-\pi, \pi]$
2. Choose latitude ϕ from $[-\frac{\pi}{2}, \frac{\pi}{2}]$ with density $\frac{1}{2} \cos \phi$.

? What is the variable's conditional distribution on a great circle?

Suppose that a random variable has a uniform distribution on a unit sphere.

? What is the variable's conditional distribution on a great circle?

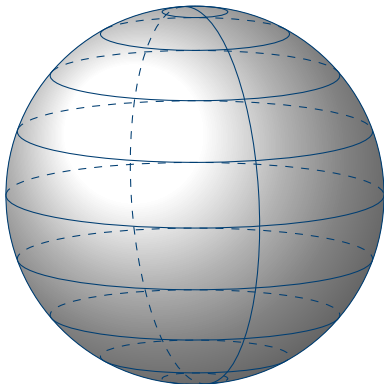


For an equator
(latitude $\phi = 0$)

$$f(\lambda|\phi = 0) = \frac{1}{2\pi}$$

Suppose that a random variable has a uniform distribution on a unit sphere.

? What is the variable's conditional distribution on a great circle?



For a line of longitude
(with $\lambda = 0$)

$$f(\phi|\lambda = 0) = \frac{1}{2} \cos \phi$$

$$f(\lambda|\phi = 0) = \frac{1}{2\pi}$$

$$f(\phi|\lambda = 0) = \frac{1}{2} \cos \phi$$

One is uniform on the circle, while the other is not!
Yet both seem to be referring to the same great circle in different coordinate systems...

Many quite futile arguments have raged – between otherwise competent probabilists – over which of these results is ‘correct’.

–E.T. Jaynes

Suppose we are given a bivariate normal distribution with mean 0 and precision $\Sigma^{-1} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$:

$$f([x, y]^T) = \frac{1}{2\pi\sqrt{|\Sigma|}} e^{-\frac{1}{2}[x, y]\Sigma^{-1}[x, y]^T} = \frac{\sqrt{1-\rho^2}}{2\pi} e^{-\frac{1}{2}(x^2+y^2-2\rho xy)}$$

Its marginal pdfs are,

$$f(x) = \sqrt{\frac{1-\rho^2}{2\pi}} e^{-\frac{1}{2}(1-\rho^2)x^2}, \quad f(y) = \sqrt{\frac{1-\rho^2}{2\pi}} e^{-\frac{1}{2}(1-\rho^2)y^2}$$

? What is the probability of x conditioned on $y = 0$?

$$f([x, y]^T) = \frac{\sqrt{1 - \rho^2}}{2\pi} e^{-\frac{1}{2}(x^2 + y^2 - 2\rho xy)}$$

? What is the probability of x conditioned on $y = 0$?

That's easy!

$$\begin{aligned} f(x|y = y_0) &= \frac{f(x, y = y_0)}{f(y = y_0)} = \frac{\frac{\sqrt{1 - \rho^2}}{2\pi} e^{-\frac{1}{2}(x^2 + y_0^2 - 2\rho xy_0)}}{\sqrt{\frac{1 - \rho^2}{2\pi}} e^{-\frac{1}{2}(1 - \rho^2)y_0^2}} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x - \rho y_0)^2} \end{aligned}$$

so,

$$f(x|y = 0) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

Now, let's perform a change of variables:

Set $u = \frac{y}{g(x)}$ for some $g(x)$ with $0 < g(x) < \infty$. The Jacobian is given by $\frac{\partial(x,u)}{\partial(x,y)} = \frac{1}{g(x)}$, so the joint pdf is given by,

$$f(x, u) = \frac{\sqrt{1 - \rho^2}}{2\pi} e^{-\frac{1}{2}(x^2 + u^2 g^2(x) - 2\rho x u g(x))} g(x)$$

? What is the probability of x conditioned on $u = 0$?

$$f(x, u) = \frac{\sqrt{1 - \rho^2}}{2\pi} e^{-\frac{1}{2}(x^2 + u^2 g^2(x) - 2\rho x u g(x))} g(x)$$

? What is the probability of x conditioned on $u = 0$?

That's easy as well!

$$\begin{aligned} f(x|u = u_0) &= \frac{f(x, u = u_0)}{f(u = u_0)} = \frac{\frac{\sqrt{1 - \rho^2}}{2\pi} e^{-\frac{1}{2}(x^2 + u_0^2 g^2(x) - 2\rho x u_0 g(x))} g(x)}{\sqrt{\frac{1 - \rho^2}{2\pi}} e^{-\frac{1}{2}(1 - \rho^2) u_0^2 g^2(x)}} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x - \rho u_0 g(x))^2} g(x) \end{aligned}$$

so,

$$f(x|u = 0) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} g(x)$$

$$u = \frac{y}{g(x)}, g(x) \text{ arbitrary}$$

Define event A as ' $y = 0$ ', and define event B as ' $u = 0$ ';
 $y = 0$ if and only if $u = 0$, so $A = B$.

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

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

- ▶ We conditioned on the same event and obtained different results!

- ▶ The problem boils down to the way in which we define a probability conditioned on a zero-measure set:

$$f(x|y=0) = \frac{f(x, y=0)}{f(y=0)} = \mathbf{lim}_{\epsilon \rightarrow 0} \frac{f(x, -\epsilon < y < \epsilon)}{f(-\epsilon < y < \epsilon)}.$$

- ▶ This limit is different from the limit for u :

$$\begin{aligned} f(x|u=0) &= \frac{f(x, u=0)}{f(u=0)} = \lim_{\epsilon \rightarrow 0} \frac{f(x, -\epsilon < u < \epsilon)}{f(-\epsilon < u < \epsilon)} \\ &= \lim_{\epsilon \rightarrow 0} \frac{f(x, -\epsilon < \frac{y}{g(x)} < \epsilon)}{f(-\epsilon < \frac{y}{g(x)} < \epsilon)} = \mathbf{lim}_{\epsilon \rightarrow 0} \frac{f(x, -g(x)\epsilon < y < g(x)\epsilon)}{f(-g(x)\epsilon < y < g(x)\epsilon)}. \end{aligned}$$

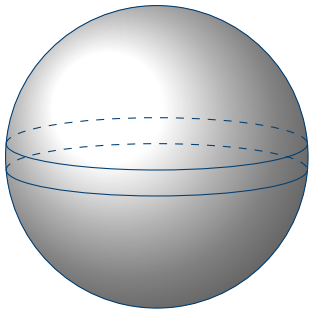
$$\lim_{\epsilon \rightarrow 0} \frac{f(x, -\epsilon < y < \epsilon)}{f(-\epsilon < y < \epsilon)}$$

$$\lim_{\epsilon \rightarrow 0} \frac{f(x, -g(x)\epsilon < y < g(x)\epsilon)}{f(-g(x)\epsilon < y < g(x)\epsilon)}$$

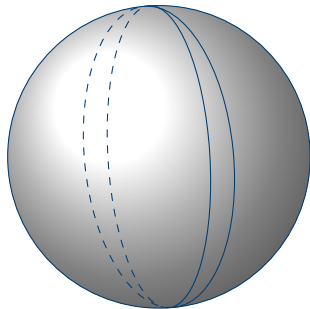
- ▶ The first limit constrains y to successively narrower horizontal strips $[-\epsilon, \epsilon]$. But the proposition “ $y = 0$ ” could also be defined using the sequence $[-g(x)\epsilon, g(x)\epsilon]$ which changes with x .
- ▶ Specifying “ $y = 0$ ” without any qualifications is ambiguous; it tells us to pass to a measure zero limit, but not which of any number of limits is intended.

Going back to the Borel–Kolmogorov paradox...

$$f(\lambda | -\epsilon < \phi < \epsilon) \rightarrow \frac{1}{2\pi}$$

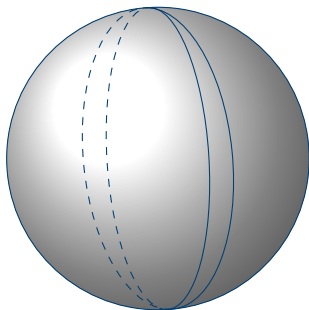
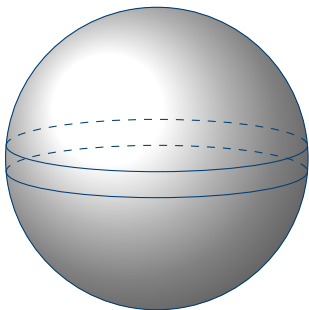


$$f(\phi | -\epsilon < \lambda < \epsilon) \rightarrow \frac{1}{2} \cos \phi$$



... the term 'great circle' is ambiguous until we specify what limiting operation is to produce it. The intuitive symmetry argument presupposes the equatorial limit; yet one eating slices of an orange might presuppose the other.

–E.T. Jaynes



- ▶ Jaynes, E.T. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press. pp. 467470. ISBN 0-521-59271-2. MR 1992316.

Chapter 15: PARADOXES OF PROBABILITY THEORY