Bishop Questions Solutions

Ex (1.1)

We have been given the following identities:

$$\int_{-\infty}^{\infty} \exp\left\{-\frac{\lambda}{2}x^2\right\} dx = \left(\frac{2\pi}{\lambda}\right)^{1/2} \quad \text{—— (1)}$$

and

$$\prod_{i=1}^{d} \int_{-\infty}^{\infty} e^{-x_i^2} dx_i = S_d \int_0^{\infty} e^{-r^2} r^{d-1} dr \quad \text{—— (2)}$$

where S_d is the surface area of the unit sphere in d- dimensions. We need to prove that:

$$S_d = \frac{2\pi^{d/2}}{\Gamma(d/2)} \quad \text{—— (3)}$$

where $\Gamma(x)$ is the gamma function defined by:

$$\Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u} du \quad \text{—— (4)}$$

Using (1) and substituting $\lambda = 2$, we have:

$$\int_{-\infty}^{\infty} \exp\left\{-\frac{2}{2}x^2\right\} dx = \left(\frac{2\pi}{2}\right)^{1/2}$$

or

$$\int_{-\infty}^{\infty} \exp\{-x^2\} dx = \pi^{1/2} \quad \text{—— (5)}$$

Substituting in (2),

$$\prod_{i=1}^{d} \int_{-\infty}^{\infty} e^{-x_i^2} dx_i = \prod_{i=1}^{d} \pi^{1/2}$$

$$\cancel{\Rightarrow \pi^{d/2} = S_d \int_0^{\infty} e^{-u} u^{(d-1)/2} \frac{du}{2}}$$

$$\Rightarrow \pi^{d/2} = S_d \int_0^{\infty} e^{-r^2} r^{d-1} dr$$

Now, converting the RHS into an equation in 'u' where $x^2 = u$ and $dx = \frac{du}{2\sqrt{u}}$, we get

$$\pi^{d/2} = S_d \int_0^\infty e^{-u} \, u^{(d-1)/2} \cdot \frac{du}{2u^{1/2}}$$

$$\Rightarrow 2\pi^{d/2} = S_d \int_0^\infty e^{-u} \, u^{(d/2 - 1)} \, du$$

The term inside integral on RHS reduces to the definition of $\Gamma(x)$ with $x = \frac{d}{2}$

Hence, we have

$$2\pi^{d/2} = S_d \cdot \Gamma(d/2)$$

$$\Rightarrow \boxed{S_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}}$$ Hence, proved.

Next, we have been given that $\Gamma(1) = 1$ — (A) and $\Gamma(3/2) = \frac{\sqrt{\pi}}{2}$ — (B)

We need to prove that (3) reduces to well known expressions for $d = 2$ and $d = 3$.

To start with, we know that :

[where 'r' is the radius]

$$S_2 = 2\pi r$$
$$= 2\pi \quad \text{at } r = 1 \quad \text{— (6)}$$

and
$$S_3 = 4\pi r^2$$
$$= 4\pi \quad \text{at } r = 1 \quad \text{— (7)}$$

Using (3), we have :

$$S_2 = \frac{2\pi^{2/2}}{\Gamma(2/2)} \quad [d = 2]$$

$$= \frac{2\pi}{\Gamma(1)} = 2\pi \quad \text{(from (A))}$$

Scanned by CamScanner

$$\Rightarrow \boxed{S_2 = 2\pi} \quad \text{— same as (6)}$$

Similarly, $S_3 = \dfrac{2\pi^{3/2}}{\Gamma(3/2)}$

$= \dfrac{2\pi^{3/2}}{\sqrt{\pi}/2}$ using (B)

$$\Rightarrow \boxed{S_3 = 4\pi} \quad \text{— same as (7)}$$

Hence, S_2, S_3 values reduce to correct expressions.

1.2

Using (3), we need to show that the volume of a hypersphere of radius 'a' in d-dimensions is given by

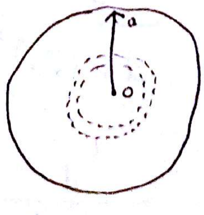
$$V_d = \frac{S_d a^d}{d} \quad \text{———— (8)}$$

We know that the surface area of any d-dimensional sphere with radius 'a' is given as:

$$S_{da} = S_d \cdot a^{d-1}$$

where $S_d = S_{d1}$, putting $a = 1$

For a d-dimensional sphere, the volume of all of the sphere may be expressed as the sum of all surface areas of spheres of radii from zero through 'a'



Thus, we can say that the volume of a sphere with d-dimensions and radius 'a' is given as:

Scanned by CamScanner

$$V_{da} = \int_0^a S_d \, a^{d-1} \, da$$

$$= S_d \int_0^a a^{d-1} \, da$$ [$\because S_d$ is for a unit sphere and hence, independent of 'a']

$$= S_d \left[\frac{a^d}{d} \right]_0^a$$

$$\Rightarrow \boxed{V_{da} = \frac{S_d \cdot a^d}{d}}$$ Hence, proved

Next, we need to show that the ratio of the volume of a hypersphere of radius 'a' to the volume of a hypercube of side '2a' (i.e. The circumscribed hypercube) is given by:

$$\frac{\text{Volume of sphere}}{\text{Volume of cube}} = \frac{\pi^{d/2}}{d \cdot 2^{d-1} \cdot \Gamma(d/2)} \quad \text{——} (9)$$

We know that the volume of a d-dimensional cube with side '2a' is given as:

$$(V_d)_{cube} = (2a)^d \quad \text{——} (10)$$

Therefore, we have:

$$\frac{V_{da}}{(V_d)_{cube}} = \frac{S_d \, a^d}{d \cdot (2a)^d}$$ (from (8) & (10))

$$= \frac{S_d}{d \cdot 2^d}$$

$$= \frac{2\pi^{d/2}}{\Gamma(d/2).\, d .\, 2^{d}} \qquad \text{(From (3))}$$

$$\Rightarrow \boxed{\frac{\text{Volume of Sphere}}{\text{Volume of Cube}} = \frac{\pi^{d/2}}{d.\, 2^{d-1}.\, \Gamma(d/2)}}$$ Hence, proved.

For the next part, we need to use Stirling's approximation:

$$\Gamma(x+1) \approx (2\pi)^{1/2} e^{-x} x^{x+1/2} \qquad \text{— (11)}$$

which is valid when x is large.

We need to prove that as $d \to \infty$, the ratio in (9) goes to zero.

Using integration on (4), we get the property

$$\Gamma(\alpha) = (\alpha-1)\,\Gamma(\alpha-1)$$

Putting $\alpha = d/_2 + 1$, we have:

$$\Gamma\left(\frac{d}{2}+1\right) = \frac{d}{2}\,\Gamma\left(\frac{d}{2}\right)$$

$$\Rightarrow \Gamma\left(\frac{d}{2}\right) = \frac{2}{d}\,\Gamma\left(\frac{d}{2}+1\right) \qquad \text{— (12)}$$

Putting x as $\frac{d}{2}$ in (11) and substituting in (12),

we get:

$$\Gamma(d/_2) = \frac{2}{d}\left[(2\pi)^{1/2} .\, e^{-d/2} \left(\frac{d}{2}\right)^{d/2 + 1/2}\right]$$

Using the above result in (9),

$$\frac{V_{da}}{(V_a)_{cube}} = \frac{\pi^{d/2}}{d.\, 2^{d-1}.\, \frac{2}{d}.\, (2\pi)^{1/2}.\, e^{-d/2}.\, \left(\frac{d}{2}\right)^{\frac{d+1}{2}}}$$

$$= \frac{\pi^{d/2}}{2^{d/2}.\, (\pi^{1/2})\, e^{-d/2}.\, d^{d/2}.\, d^{1/2}}$$

$$= \left(\frac{\pi e}{2d}\right)^{d/2} \cdot \frac{1}{(\pi d)^{1/2}}$$

Therefore, as $d \to \infty$, $\frac{1}{d} \to 0$

This gives:

$$\boxed{\frac{V_{da}}{(V_o)_{cube}} = 0 \times 0 = 0}$$ Hence, proved

Next, we need to show that the ratio of the distance from the center of the hypercube to one of the corners, divided by the perpendicular distance to one of the edges is $\sqrt{d}$ and thus goes to ∞, as $d \to \infty$

We know that the length of diagonal of d-dimensional hypercube with side '$2a$' is given as:

$$l_c = \left(\sum_{i=1}^{d} (2a)^2\right)^{1/2} \quad \text{—— (13)}$$

Therefore, length from center to one of the corners is $\frac{l_c}{2} = \frac{2a\sqrt{d}}{2}$

Also, the perpendicular to one of the edges will have length half of that of side.

$$\therefore \ l_p = a \quad \text{—— (14)}$$

Thus, $$\frac{l_c/2}{l_p} = \frac{(2a\sqrt{d})/2}{a}$$

$$= \sqrt{d}$$

Hence, as $d \to \infty$, this ratio goes to ∞.

1.3

In this question, using (8), we need to prove that the fraction of the volume of the sphere which lies at values of the radius between $a-\epsilon$ and a, where $0 < \epsilon < a$, is given as:

$$f = 1 - \left(1 - \frac{\epsilon}{a}\right)^d$$

We can write the fraction of volume of the sphere lying between spheres of radii $a-\epsilon$ and a as:

$$\frac{V_{da} - V_{d(a-\epsilon)}}{V_{da}} \qquad \text{—— (15)}$$

Therefore,

$$f = \frac{S_d \cdot \frac{a^d}{d} - S_d \frac{(a-\epsilon)^d}{d}}{\frac{S_d \cdot a^d}{d}}$$

$$= \frac{a^d - (a-\epsilon)^d}{a^d}$$

$$= 1 - \left(\frac{a-\epsilon}{a}\right)^d$$

$$\Rightarrow \boxed{f = 1 - \left(1 - \frac{\epsilon}{a}\right)^d} \quad \text{Hence, proved.}$$

Now, we need to show that for any ϵ no matter how small, this fraction tends to 1 as $d \rightarrow \infty$

We know that

$$0 < \epsilon < a \qquad \text{(given)}$$

$$\Rightarrow \quad 0 < \frac{\epsilon}{a} < 1$$

$$\Rightarrow \quad 0 < 1 - \frac{\epsilon}{a} < 1$$

Hence, as $d \to \infty$, $\left(1 - \frac{\epsilon}{a}\right)^d \to 0$

Therefore $f_{d \to \infty} = 1 - 0$

$= 1$ Hence, proved.

Now, for the next part, we have :

$$\frac{\epsilon}{a} = 0.01$$

$$\Rightarrow \quad 1 - \frac{\epsilon}{a} = 0.99$$

$$\therefore \quad f = 1 - (0.99)^d$$

$$\Rightarrow \quad f_{d=2} = 1 - (0.99)^2$$

$$= 0.0199$$

$$\Rightarrow \quad f_{d=10} = 1 - (0.99)^{10}$$

$$= 0.0956$$

$$\Rightarrow \quad f_{d=1000} = 1 - (0.99)^{1000}$$

$$= 0.949957$$

Next, we need to evaluate the fraction of the volume of the sphere which lies inside the radius $\frac{a}{2}$ for $d = 2, 10, 1000$.

Using (8), we write the ratio R as

$$R = \frac{\frac{S_d \cdot (a/2)^d}{d}}{\frac{S_d \cdot (a)^d}{d}} = \frac{a^d}{2^d \cdot a^d} = \frac{1}{2^d}$$

$$\Rightarrow \quad R_{d=2} = \frac{1}{2^2} = 0.25$$

$$R_{d=10} = \frac{1}{2^{10}} = 0.000976$$

$$R_{d=1000} = \frac{1}{2^{1000}} \approx 0$$

Thus, we infer that almost no point is inside the radius $\frac{a}{2}$ as d increases.

———X———

1.4 In this question, we are given a probability density function $p(x)$ in d-dimensions, which is a function of radius $r = ||x||$ and which has a Gaussian form

$$p(x) = \frac{1}{(2\pi\sigma^2)^{d/2}} \exp\left(-\frac{||x||^2}{2\sigma^2}\right) \quad \text{——— (16)}$$

We need to prove that by changing variables from Cartesian to polar coordinates, the probability mass inside a thin shell of radius r and thickness ϵ is given as $p(r)\epsilon$ where

$$p(r) = \frac{S_d r^{d-1}}{(2\pi\sigma^2)^{d/2}} \exp\left(-\frac{r^2}{2\sigma^2}\right) \quad \text{——— (17)}$$

Then, we need to show that the function $p(r)$ has a single maximum which, for large values of d, is located at $\hat{r} \approx \sqrt{d}\,\sigma$. Finally, by considering $p(\hat{r}+\epsilon)$ where $\epsilon \ll \hat{r}$ we need to show that for large d

$$p(\hat{r}+\epsilon) = p(\hat{r}) \exp\left(-\frac{\epsilon^2}{\sigma^2}\right) \quad \text{——— (18)}$$

We have, $r = ||x|| = \left(\sum_{i=1}^{d} x_i^2\right)^{1/2}$

Therefore,

$$p(x) = \frac{1}{(2\pi\sigma^2)^{d/2}} \exp\left(-\frac{||x||^2}{2\sigma^2}\right)$$

$$= \frac{1}{(2\pi\sigma^2)^{d/2}} \exp\left(-\frac{\sum_{i=1}^{d} x_i^2}{2\sigma^2}\right)$$

Now, probability mass inside a shell of radius r and thickness ϵ is $= \int_{||x||=r}^{||x||=r+\epsilon} p(\vec{x})\, d\vec{x}$

$$= \iint \cdots \int_{||x||=r}^{||x||=r+\epsilon} p(x_1, x_2 \cdots x_d)\, dx_1\, dx_2 \cdots dx_d$$

$$= \frac{1}{(2\pi\sigma^2)^{d/2}} \iint \cdots \int_{||x||=r}^{||x||=r+\epsilon} e^{-x_1^2/2\sigma^2} \cdot e^{-x_2^2/2\sigma^2} \cdots e^{-x_d^2/2\sigma^2} dx_1 \ldots dx_d$$

$$= \frac{1}{(2\pi\sigma^2)^{d/2}} \prod_{i=1}^{d} \int_{||x||=r}^{||x||=r+\epsilon} (2\sigma^2)\, e^{-x_i^2/2\sigma^2} \cdot \frac{dx_i}{2\sigma^2}$$

$$= \frac{(2\sigma^2)^d}{(2\pi\sigma^2)^{d/2}} \cdot S_d \cdot \int_{r}^{r+\epsilon} e^{-r^2/2\sigma^2} \left(\frac{r}{2\sigma^2}\right)^{d-1} \cdot \frac{dr}{2\sigma^2}$$

(From ② in 1.1)

$$= \frac{1}{(2\pi\sigma^2)^{d/2}} \cdot S_d \int_{r}^{r+\epsilon} e^{-r^2/2\sigma^2} \cdot r^{d-1}\, dr$$

Since $dr \approx \epsilon$, we can ignore the integral

$$\approx \frac{1}{(2\pi\sigma^2)^{d/2}} \cdot r^{d-1} \cdot e^{-r^2/2\sigma^2} \cdot S_d\, dr$$

$$= \frac{S_d\, r^{d-1}\, e^{-r^2/2\sigma^2}}{(2\pi\sigma^2)^{d/2}} \cdot \epsilon = p(r)\, \epsilon$$

Hence, proved.

For the second part, we differentiate (17) w.r.t. r partially

$$\frac{\delta p(r)}{\delta r} = \left[\frac{(d-1)\, r^{d-2}\, S_d}{(2\pi\sigma^2)^{d/2}} - \frac{2r}{2\sigma^2} \cdot \frac{S_d \cdot r^{d-1}}{(2\pi\sigma^2)^{d/2}} \right] e^{-r^2/2\sigma^2}$$

$$= 0 \quad (\text{for optimum value})$$

$$\Rightarrow \left[(d-1)\, r^{d-2} - \frac{r^d}{\sigma^2} \right] \cdot \frac{S_d}{(2\pi\sigma^2)^{d/2}} \; e^{-r^2/2\sigma^2} = 0$$

The second term cannot be zero.

$$\Rightarrow (d-1)\, r^{d-2} - \frac{r^d}{\sigma^2} = 0$$

$$\Rightarrow r^{d-2}\left((d-1)\sigma^2 - r^2 \right) = 0$$

$$\Rightarrow r^2 = \sigma^2 (d-1)$$

$$\Rightarrow r \approx \sqrt{d} \cdot \sigma \quad \text{for large } d \qquad \text{—— (19)}$$

For $r < \sqrt{d} \cdot \sigma$, we have

$$\text{sign}\left(\frac{\delta p(r)}{\delta r} \right) = \frac{|d\sigma^2 - r^2|}{d\sigma^2 - r^2} \rightarrow +ve \qquad \text{—— (20)}$$

For $r > \sqrt{d} \cdot \sigma$, we have

$$\text{sign}\left(\frac{\delta p(r)}{\delta r} \right) = -ve \qquad \text{—— (21)}$$

From (19), (20) and (21), we can say that $\sqrt{d} \cdot \sigma$ is a maxima and is unique.

Hence, proved.

Scanned by CamScanner

Now, for the final part, we have

$$p(r+\epsilon) = \frac{S_d (r+\epsilon)^{d-1}}{(2\pi\sigma^2)^{d/2}} \cdot \exp\left\{-\frac{(r+\epsilon)^2}{2\sigma^2}\right\}$$

Therefore,

$$\frac{p(\hat{r}+\epsilon)}{p(\hat{r})} = \frac{(\hat{r}+\epsilon)^{d-1}}{(\hat{r})^{d-1}} \cdot \frac{\exp\left\{-\frac{(\hat{r}+\epsilon)^2}{2\sigma^2}\right\}}{\exp\left\{\frac{-\hat{r}^2}{2\sigma^2}\right\}}$$

$$= \left[1+\frac{\epsilon}{\hat{r}}\right]^{d-1} \exp\left\{\frac{\hat{r}^2}{2\sigma^2} - \frac{(\hat{r}+\epsilon)^2}{2\sigma^2}\right]$$

$$= \exp\left\{\ln\left(1+\frac{\epsilon}{\hat{r}}\right)^{d-1}\right\} \exp\left\{\frac{1}{2\sigma^2}(\hat{r}^2-(\hat{r}^2+\epsilon^2+2\hat{r}\epsilon))\right\}$$

$$= \exp\left\{(d-1)\ln\left(1+\frac{\epsilon}{\hat{r}}\right) - \frac{\epsilon^2}{2\sigma^2} - \frac{\epsilon\hat{r}}{\sigma^2}\right\} \quad (22)$$

Now, as $d \to \infty$, the higher order terms of (22)'s Taylor Series go to zero. This is because $\hat{r} \propto \sqrt{d}$.

Taking first two terms from Taylor series, we get:

$$\frac{p(\hat{r}+\epsilon)}{p(\hat{r})} = \exp\left[(d-1)\left[\frac{\epsilon}{\hat{r}} - \frac{\epsilon^2}{2\hat{r}^2}\right] - \frac{\epsilon^2}{2\sigma^2} - \frac{\epsilon\hat{r}}{\sigma^2}\right]$$

Substituting $\hat{r}$ with $\sigma\sqrt{d-1}$,

$$\frac{p(\hat{r}+\epsilon)}{p(\hat{r})} = \exp\left\{\frac{\epsilon\sqrt{d-1}}{\sigma} - \frac{\epsilon^2}{2\sigma^2} - \frac{\epsilon^2}{2\sigma^2} - \frac{\epsilon\sqrt{d-1}}{\sigma}\right\}$$

$$= \exp\left\{-\frac{\epsilon^2}{\sigma^2}\right\}.$$

Hence, proved.