Ex. 2.4

Consider set of training data drawn from Spherical Multinormal dist. $X \sim N(0, \mathbb{I}_P)$

First we will find the expected distance from any point to the origin.
Notice that as cov is $\mathbb{I}_P$ each dimension is $\overset{iid}{\sim} N(0,1)$ for any observation. e.g for observation $Z \sim N(0, \mathbb{I}_P)$, we have $Z_1, Z_2, \ldots, Z_P \overset{iid}{\sim} N(0,1)$.
Additionally the squared distance $Z_1^2 + \ldots + Z_P^2 \sim \chi^2_P$ from the definition of the chi squared distribution.
$\Rightarrow \mathbb{E}($ ~~?~~ squared dist from Z to origin$) = \mathbb{E}(Z_1 + \ldots + Z_P) = P$

Now suppose we take a test point x_0 from the same distribution. Clearly the expected squared distance between this point and the origin is P too.

Finally we ~~?~~ will project all training points in the direction of x_0 and calculate their expected distance from the origin.

$a = \dfrac{x_0}{\|x_0\|}$ is the associated unit vector of x_0, so $Z_i = a^T x_i$ is the projection of each training point in this direction.

Since x_i is Normal and a is constant, Z_i is normal.
$E(Z_i) = E(a^T x_i) = a^T(0) = 0$
$Var(Z_i) = Var(a^T x_i) = a^T(1)a = \dfrac{x_0^T}{\|x_0\|}\dfrac{x_0}{\|x_0\|} = 1$

$\Rightarrow Z_i \sim N(0,1)$

$Z \sim \chi^2_1$

$E(\text{Squared dist. to origin}) = E(Z^2)$ ~~?~~ $= 1$

$\Rightarrow$ For $P=10$ train points are $\sqrt{1} = 1$ s.d from origin while test point is $\sqrt{10} = 3.2$ s.d from origin.