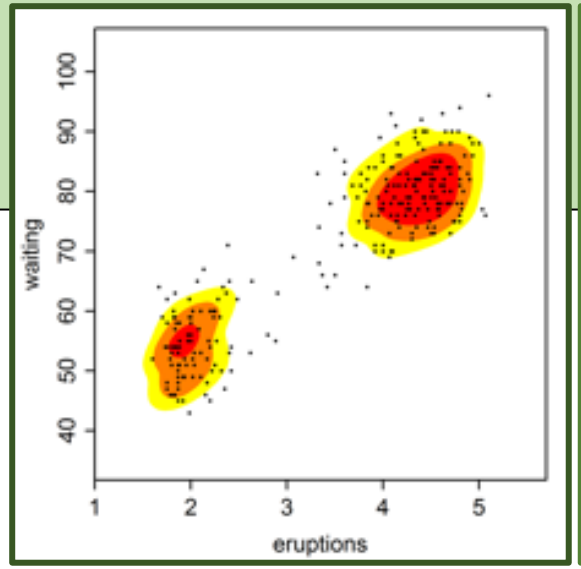
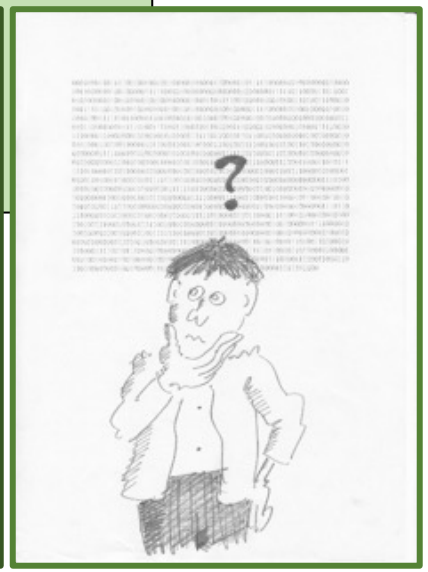
Multivar Gaussians and GMMs

Statistics and data analysis

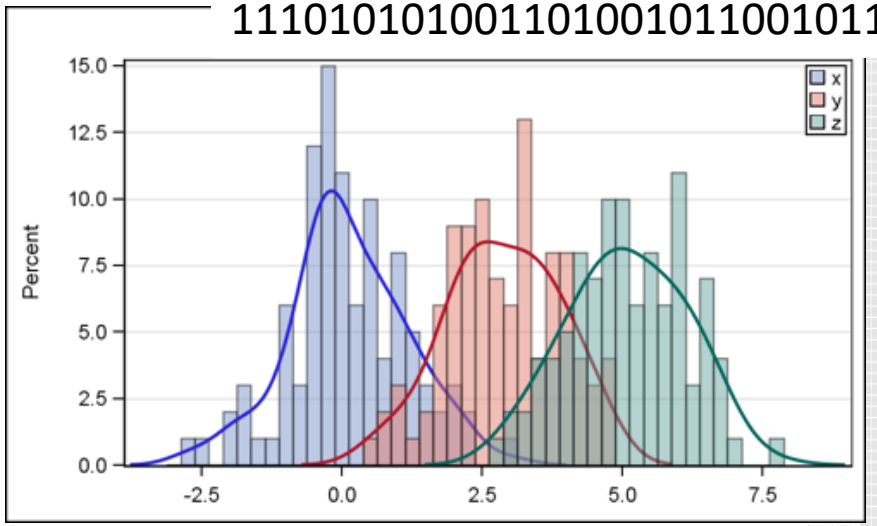
Ben Galili
Leon Anavy
Zohar Yakhini

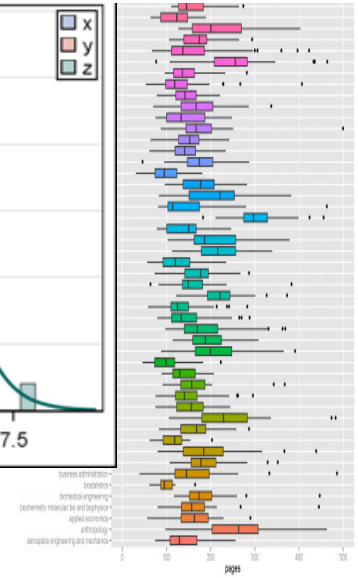
IDC, Herzeliya





0010011101010100101010100100100010
1010100010101111101011010011001001
1110101010011010010110010110110010





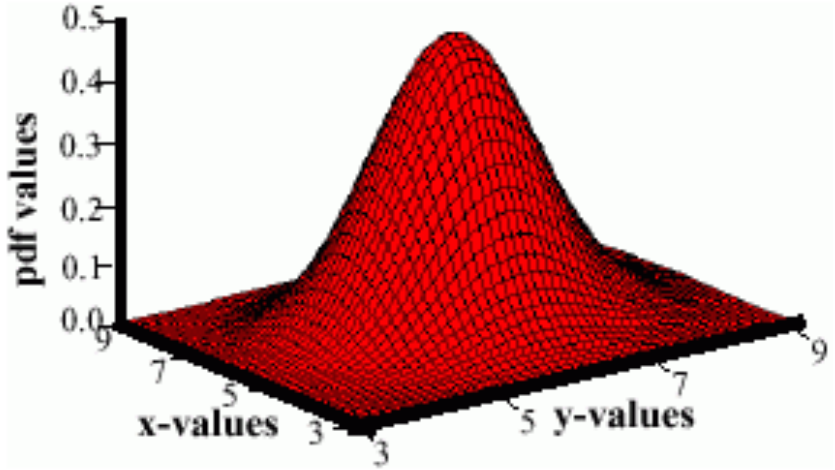
Multivariate Normal Distributions

- **A multivariate normal distribution is defined by its pdf:**

$$p(\vec{x}) = \frac{1}{\sqrt{(2\pi)^d \cdot Det\ \Sigma}}\ exp\left(-\frac{1}{2} \cdot \langle \vec{x} - \vec{\mu}, \Sigma^{-1}(\vec{x} - \vec{\mu}) \rangle\right)$$

where μ represents the mean (vector) and ∑ represents the covariance matrix.

- **The covariance is always symmetric and positive semidefinite. (invertible, Det > 0)**
- **Does the covariance matrix ∑ represent covariances?**
- **How does the shape vary as a function of the covariance? – will be discussed**



Simple case – diagonal Σ

$$
\begin{aligned}
p(x;\mu,\Sigma) &= \frac{1}{2\pi\begin{vmatrix}\sigma_1^2 & 0\\ 0 & \sigma_2^2\end{vmatrix}^{1/2}} \exp\left(-\frac{1}{2}\begin{bmatrix}x_1-\mu_1\\ x_2-\mu_2\end{bmatrix}^T\begin{bmatrix}\sigma_1^2 & 0\\ 0 & \sigma_2^2\end{bmatrix}^{-1}\begin{bmatrix}x_1-\mu_1\\ x_2-\mu_2\end{bmatrix}\right)\\
&= \frac{1}{2\pi(\sigma_1^2\cdot\sigma_2^2-0\cdot 0)^{1/2}} \exp\left(-\frac{1}{2}\begin{bmatrix}x_1-\mu_1\\ x_2-\mu_2\end{bmatrix}^T\begin{bmatrix}\frac{1}{\sigma_1^2} & 0\\ 0 & \frac{1}{\sigma_2^2}\end{bmatrix}\begin{bmatrix}x_1-\mu_1\\ x_2-\mu_2\end{bmatrix}\right)\\
&= \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{1}{2\sigma_1^2}(x_1-\mu_1)^2-\frac{1}{2\sigma_2^2}(x_2-\mu_2)^2\right)\\
&= \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{1}{2\sigma_1^2}(x_1-\mu_1)^2\right)\cdot\frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{1}{2\sigma_2^2}(x_2-\mu_2)^2\right).
\end{aligned}
$$

An alternative definition

Consider a standard bivariate normal random variable (Z_1, Z_2).
Note that we can directly define this by defining the joint pdf:

$$\varphi(x,y) = \frac{1}{2\pi} exp\left(-\frac{1}{2}(x^2+y^2)\right)$$

Using a set of parameters $\alpha, \beta, \sigma, \tau$ and ρ we now define new random variables as follows:

$$A = \sigma Z_1 + \alpha \quad \text{and} \quad B = \tau\left(\rho Z_1 + \sqrt{1-\rho^2}\cdot Z_2\right) + \beta$$

IDC HERZLIYA

Bivar normal – cont …

We can show (change of variables and some heavy calculus) that the pdf of (A, B) will have the form:

$$p(x,y) = \frac{1}{2\pi\sqrt{Det\,\Sigma}}\, exp\left(-\frac{1}{2}\cdot\left\langle (x,y)-\vec{\mu}\ ,\ \Sigma^{-1}\left((x,y)-\vec{\mu}\right)\right\rangle\right)$$

What are $\vec{\mu}$ and Σ ?

IDC HERZLIYA

Bivar normal - cont

$$E(A) = E(\sigma Z_1 + \alpha) = \alpha$$

$$V(A) = V(\sigma Z_1 + \alpha) = \sigma^2$$

IDC HERZLIYA

Similarly for B

$$E(B) = E\left(\tau\left(\rho Z_1 + \sqrt{1-\rho^2}\cdot Z_2\right) + \beta\right) = \beta$$

$$V(B) = V\left(\tau\left(\rho Z_1 + \sqrt{1-\rho^2}\cdot Z_2\right) + \beta\right)$$

$$= \tau^2\left(\rho^2 V(Z_1) + (1-\rho^2)V(Z_2)\right)$$

$$= \tau^2$$

IDC HERZLIYA

Bivar normal – cont …

We therefore have:

$$E(A) = \alpha \quad , \quad V(A) = \sigma^2 \quad ,$$

$$E(B) = \beta \quad \text{and} \quad V(B) = \tau^2$$

In fact, our construction implies that

$$A \sim Nor(\alpha, \sigma^2) \quad \text{and} \quad B \sim Nor(\beta, \tau^2)$$

IDC HERZLIYA

A useful fact

If $X \sim N(\mu, \sigma^2)$

then

$Z = (X - \mu)/\sigma$

is standard normal

If $Z \sim N(0, 1)$

then when defining

$X = \sigma Z + \mu$

we have

$X \sim N(\mu, \sigma^2)$

IDC HERZLIYA

Covariance

$$A = \sigma Z_1 + \alpha \quad \text{and} \quad B = \tau\left(\rho Z_1 + \sqrt{1-\rho^2}\cdot Z_2\right) + \beta$$

$$Cov(A,B) = E\big((A - EA)\cdot (B - EB)\big)$$

$$= E\left(\sigma Z_1 \cdot \tau\left(\rho Z_1 + \sqrt{1-\rho^2}\cdot Z_2\right)\right)$$

$$= \sigma\tau \cdot E\left(\rho Z_1^2 + \sqrt{1-\rho^2} Z_1 Z_2\right)$$

$$= \sigma\tau \cdot \rho$$

IDC HERZLIYA

Equivalence – the interpretation

Lets use an example to see how this construction reflects our first definition:

$$p(\vec{x}) = \frac{1}{\sqrt{(2\pi)^d \cdot Det\ \Sigma}}\ exp\left(-\frac{1}{2}\cdot\langle\vec{x}-\vec{\mu}, \Sigma^{-1}(\vec{x}-\vec{\mu})\rangle\right)$$

In the bivariate case, and assuming a mean at 0:

$$p(x,y) = \frac{1}{2\pi\sqrt{Det\ \Sigma}}\ exp\left(-\frac{1}{2}\cdot\langle(x,y), \Sigma^{-1}(x,y)\rangle\right)$$

IDC HERZLIYA

What is Σ ?

We constructed a bivariate random variable that has normal marginal by defining:

$$A = \sigma Z_1 \quad (+\alpha)$$

$$B = \tau\left(\rho Z_1 + \sqrt{1-\rho^2}\cdot Z_2\right) \quad (+\beta)$$

It can be shown, by a change of variables and setting α and β to 0, that the density function for our constructed bivariate normal is given by:

$$p(x,y) = \frac{1}{2\pi\sqrt{Det\,\Sigma}}\, exp\left(-\frac{1}{2}\cdot\langle(x,y),\Sigma^{-1}(x,y)\rangle\right)$$

With Σ being:

$$\Sigma = \begin{pmatrix} \sigma^2 & \sigma\tau\cdot\rho \\ \sigma\tau\cdot\rho & \tau^2 \end{pmatrix}$$

IDC HERZLIYA

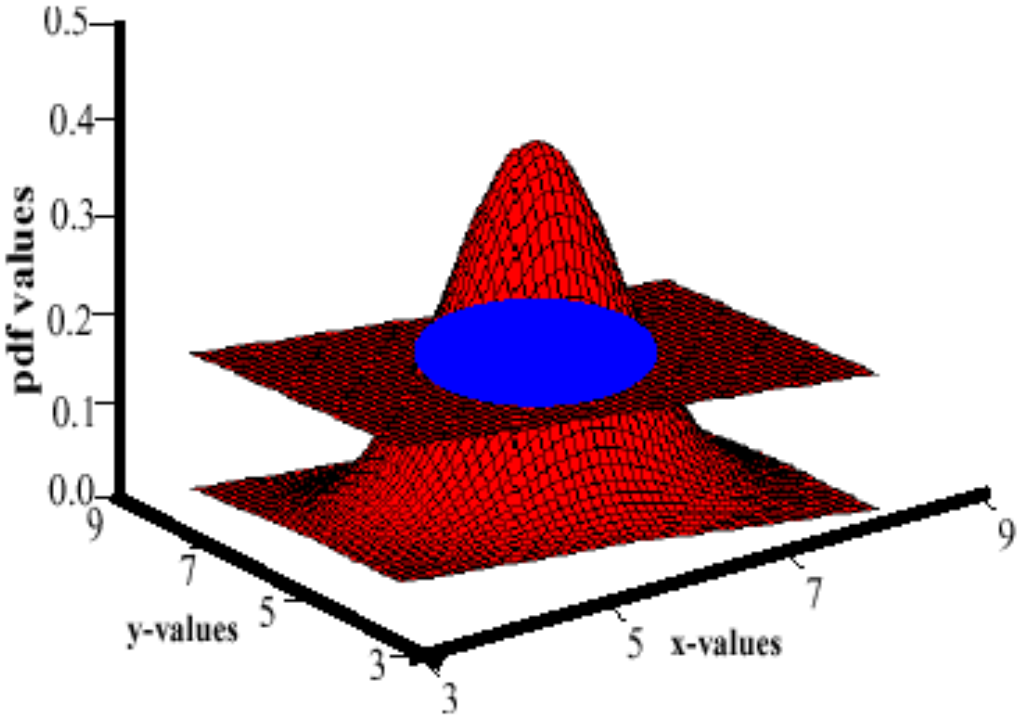
A third alternative definition

- X is multivariate normal if for any constant vector v, the random variable $\mathrm{W} = (v\,, X)$ has a univariate normal distribution.
- X is multivariate normal with a diagonal Σ iff for any constant vector v, the random variable $\mathrm{W} = (v\,, X)$ has a univariate normal distribution with $\sigma^2 = (v.^2\,, \mathrm{diag}(\Sigma))$
- In general we have $E(W) = (v, E(X))$ and $Var(W) = (v, \Sigma(v))$

IDC HERZLIYA

Support Regions or Isocontours

- Horizontal planes cut the pdf graph defining the shapes of ellipses of points that all have equal probability density.
- The shape of these support regions is determined by the covariance matrix.



IDC HERZLIYA

The diagonal case

What is the shape of isocontours for a diagonal covariance matrix?
Write:

$$c = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{1}{2\sigma_1^2}(x_1-\mu_1)^2 - \frac{1}{2\sigma_2^2}(x_2-\mu_2)^2\right)$$

$$2\pi c\sigma_1\sigma_2 = \exp\left(-\frac{1}{2\sigma_1^2}(x_1-\mu_1)^2 - \frac{1}{2\sigma_2^2}(x_2-\mu_2)^2\right)$$

$$\log(2\pi c\sigma_1\sigma_2) = -\frac{1}{2\sigma_1^2}(x_1-\mu_1)^2 - \frac{1}{2\sigma_2^2}(x_2-\mu_2)^2$$

$$\log\left(\frac{1}{2\pi c\sigma_1\sigma_2}\right) = \frac{1}{2\sigma_1^2}(x_1-\mu_1)^2 + \frac{1}{2\sigma_2^2}(x_2-\mu_2)^2$$

$$1 = \frac{(x_1-\mu_1)^2}{2\sigma_1^2\log\left(\frac{1}{2\pi c\sigma_1\sigma_2}\right)} + \frac{(x_2-\mu_2)^2}{2\sigma_2^2\log\left(\frac{1}{2\pi c\sigma_1\sigma_2}\right)}.$$

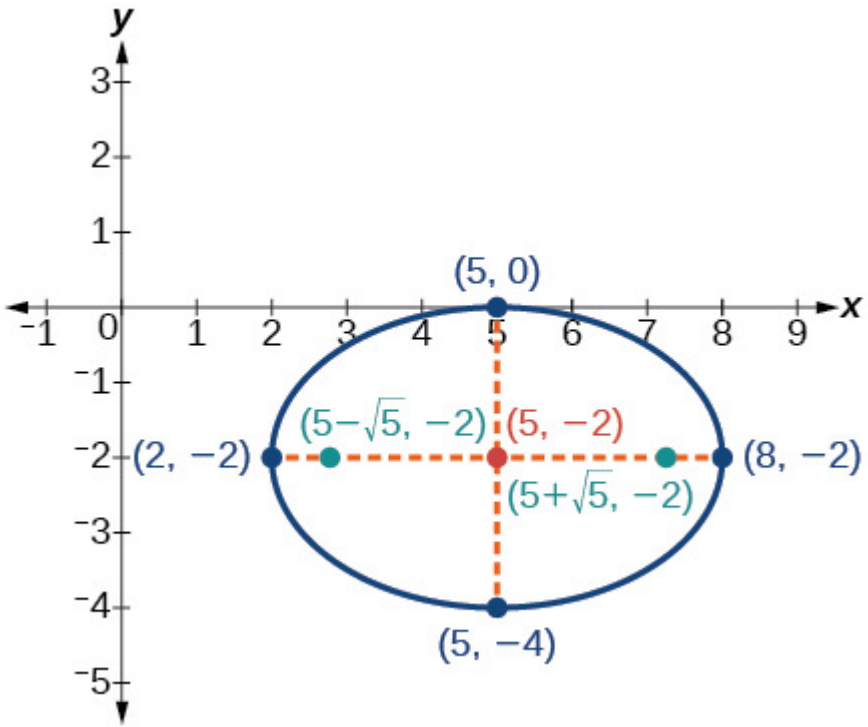
IDC HERZLIYA

Isocontours - continued

Setting

$$r_1 = \sqrt{2\sigma_1^2 \ln\left(\frac{1}{2\pi c \sigma_1 \sigma_2}\right)}$$

$$r_2 = \sqrt{2\sigma_2^2 \ln\left(\frac{1}{2\pi c \sigma_1 \sigma_2}\right)}$$

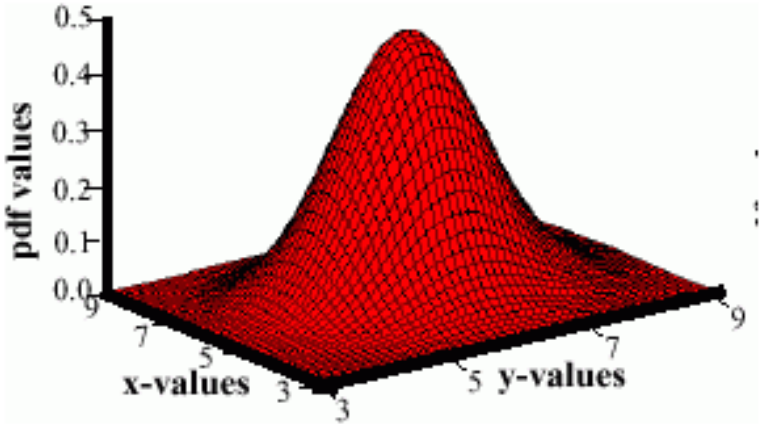


We get that (x_1, x_2) on the contour satisfies $\left(\frac{x_1-\mu_1}{r_1}\right)^2 + \left(\frac{x_2-\mu_2}{r_2}\right)^2 = 1$

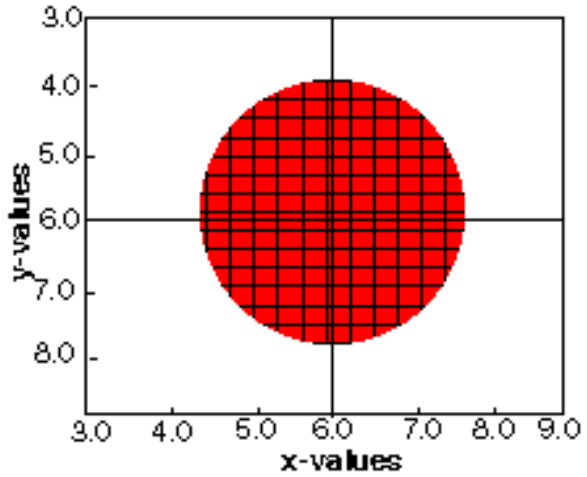
IDC HERZLIYA

Identity Covariance

Gaussian distribution



Support Region



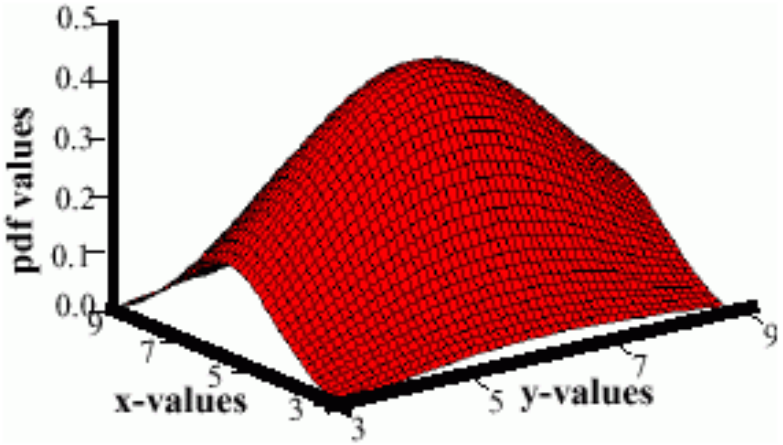
- Gaussian distribution with identity covariance matrix has equal variances in all directions

- Support region for a Gaussian distribution with identity covariance matrix $C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is a circle

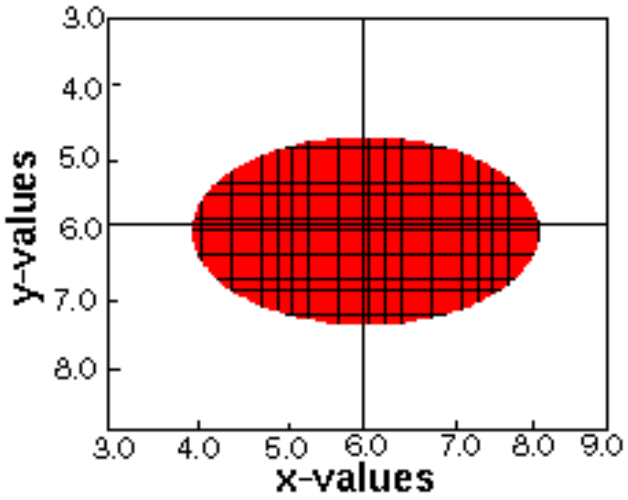
IDC HERZLIYA

Unequal Variances

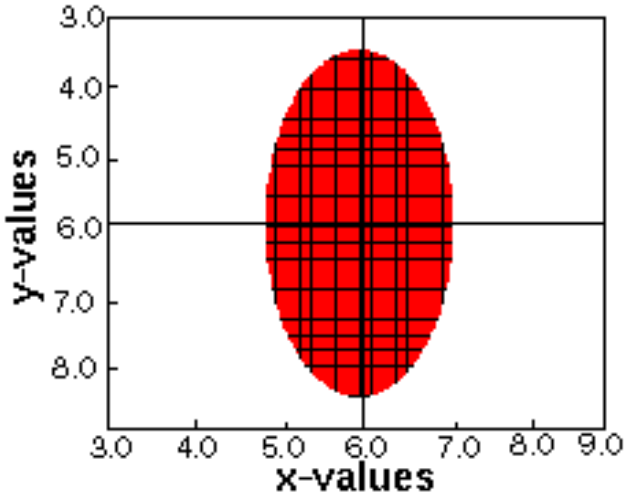
Gaussian distribution



Support Region



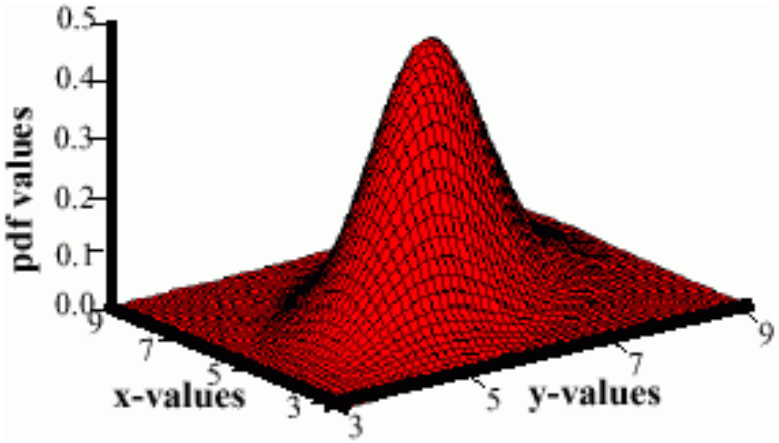
Support region for a Gaussian distribution with covariance matrix $C = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$ is an ellipse aligned with respect to the original axes



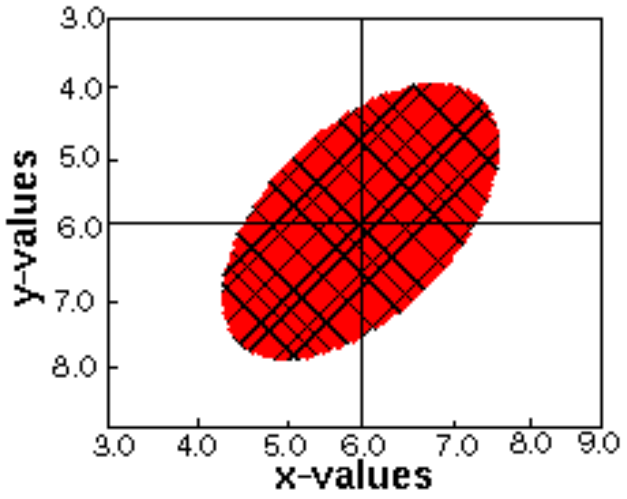
IDC HERZLIYA

Nonzero Off-Diagonal Elements

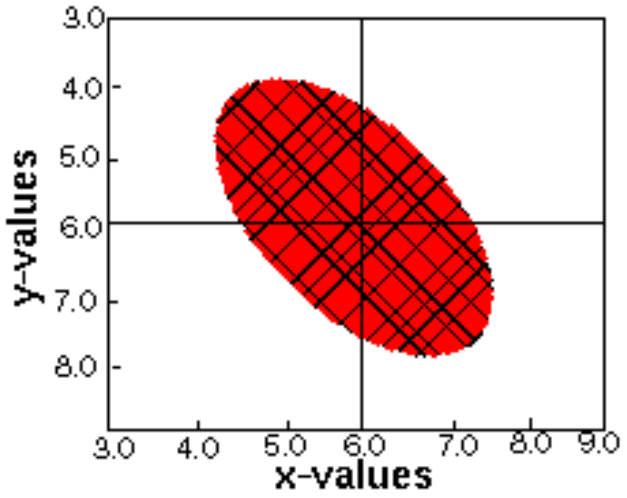
Gaussian distribution



Support Region



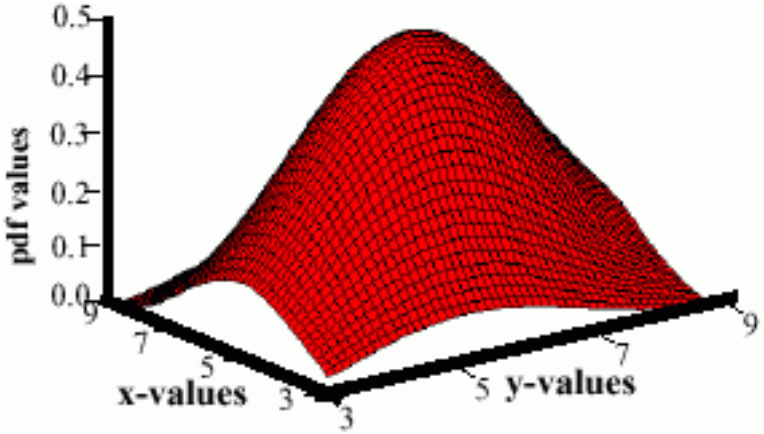
Support region for a Gaussian distribution with covariance matrix $C = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ is an ellipse rotated at an angle of 45 degree with respect to the original axes



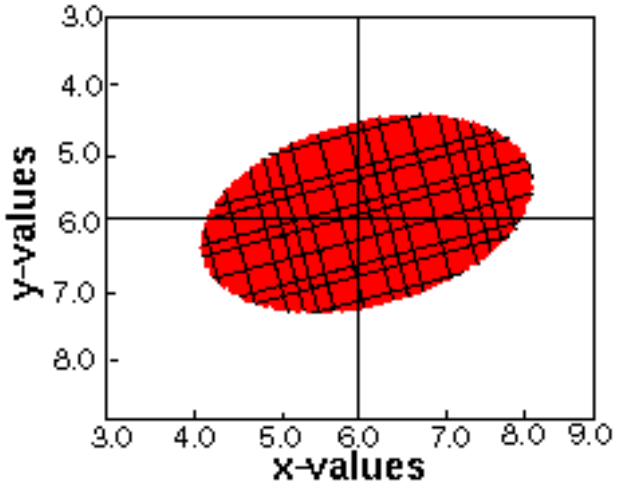
IDC HERZLIYA

Unconstrained or “Full” Covariance

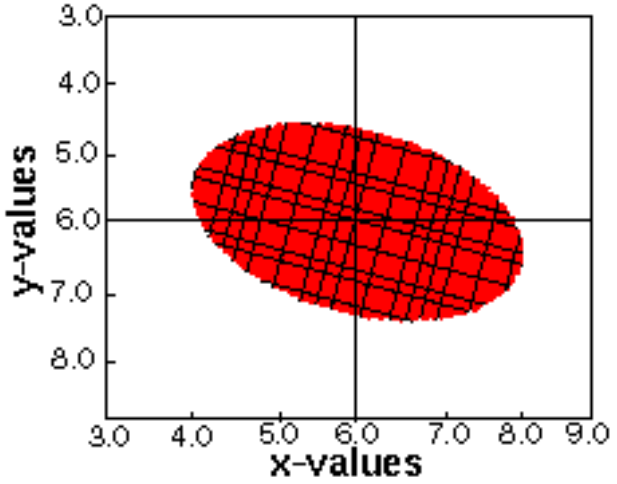
Gaussian distribution



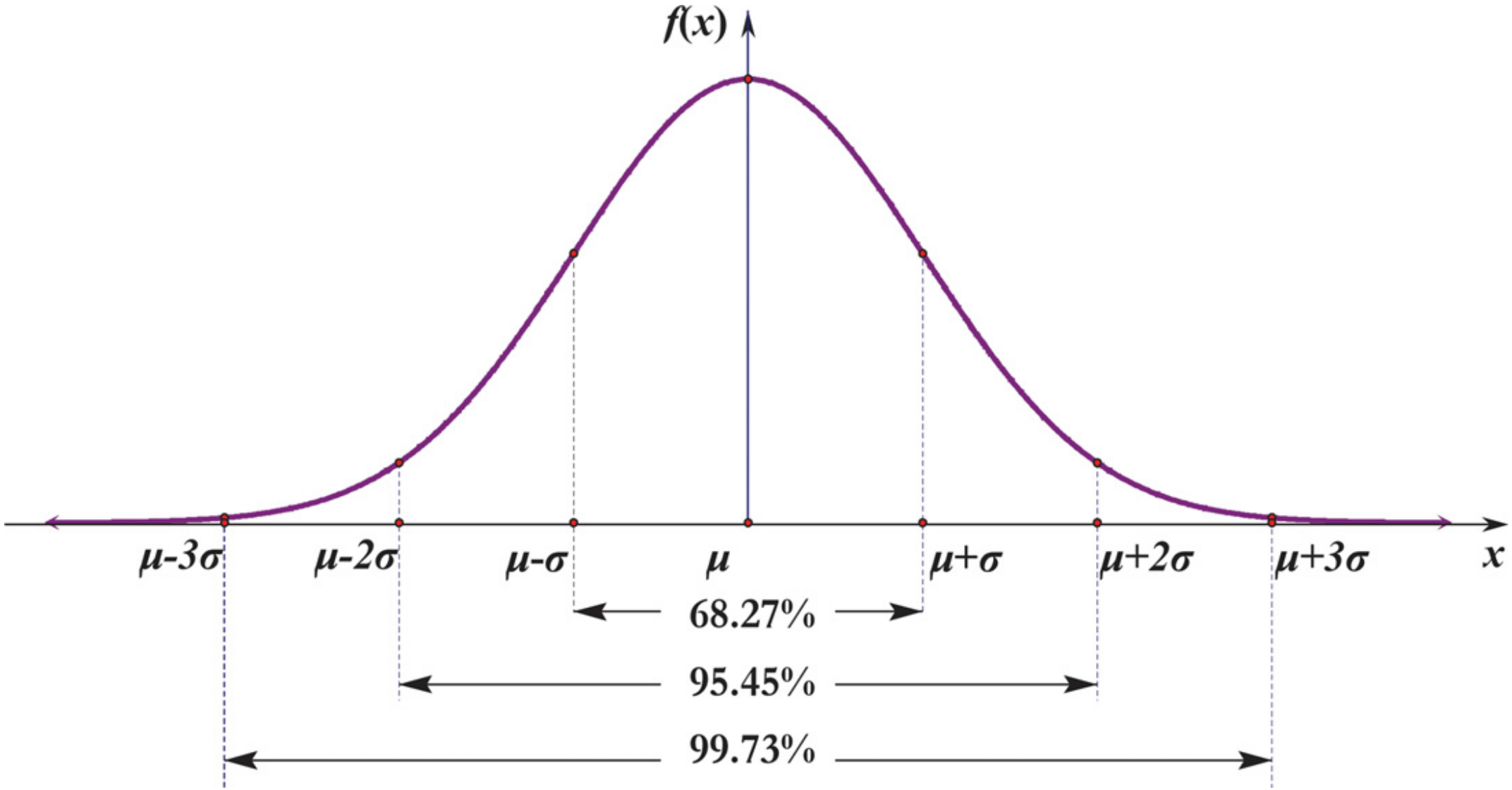
Support Region



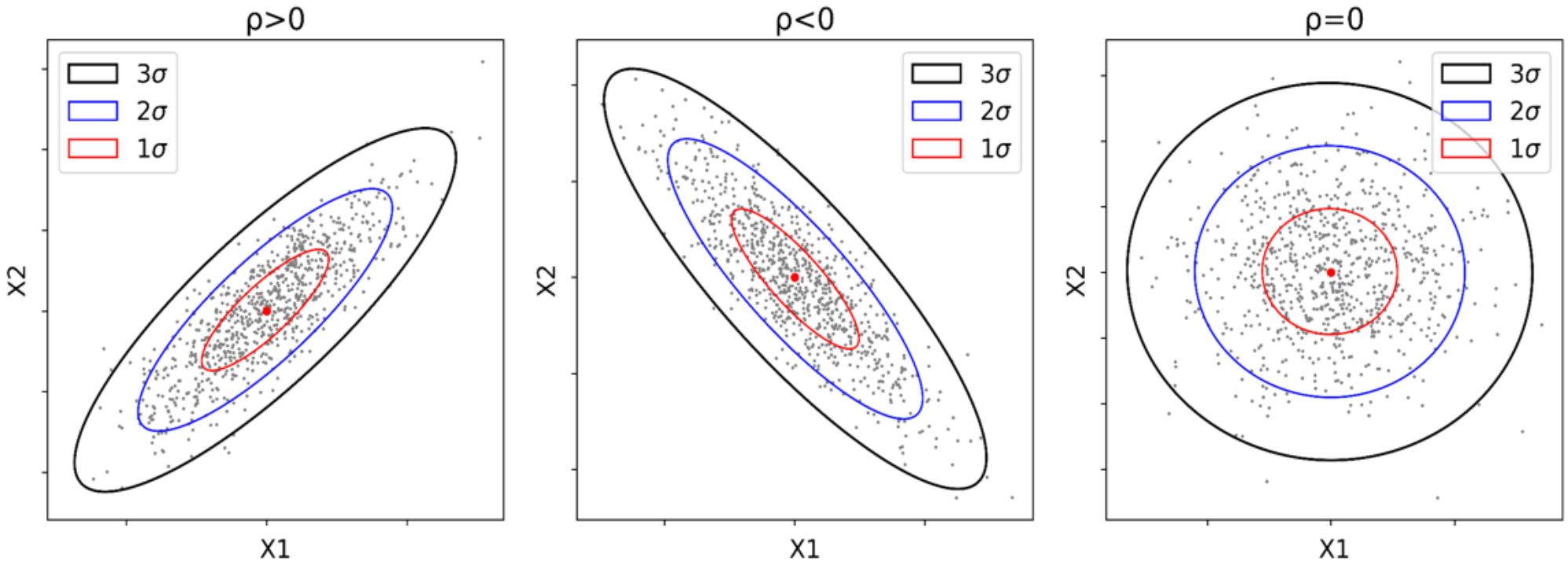
Support region for a Gaussian distribution with covariance matrix $C = \begin{bmatrix} 5 & 0.5 \\ 0.5 & 2 \end{bmatrix}$ is an ellipse rotated at an angle of 9 degrees and -81 degrees with respect to the original axes



IDC HERZLIYA



IDC HERZLIYA



IDC HERZLIYA

Table 2. Magnification ratios of scaled SDHE corresponding to different specified confidence levels with space dimensionality not exceeding 10.

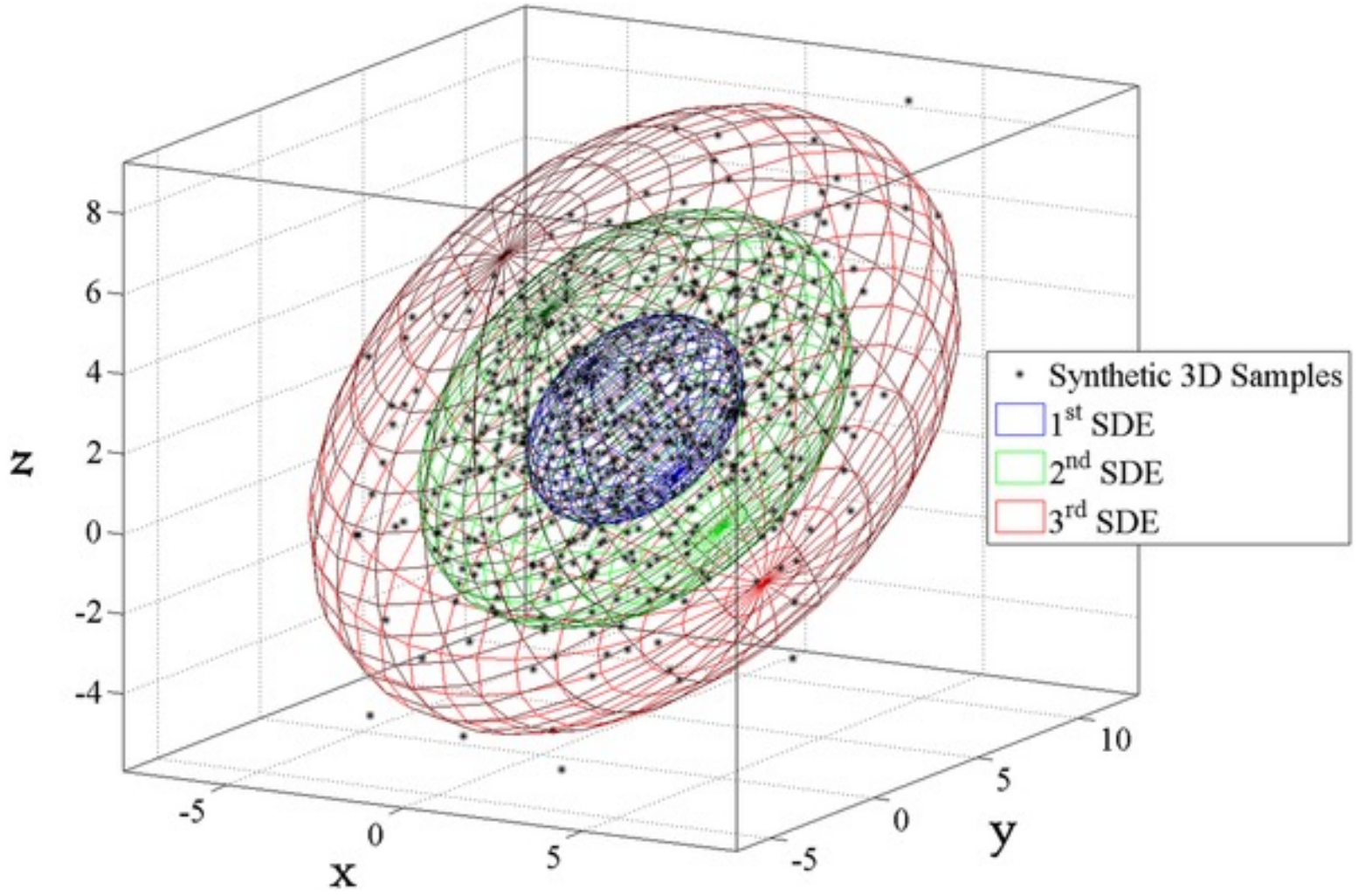
Dimensionality	Confidence Level (%)					
	80.0	85.0	90.0	95.0	99.0	99.9
1	1.2816	1.4395	1.6449	1.9600	2.5758	3.2905
2	1.7941	1.9479	2.1460	2.4477	3.0349	3.7169
3	2.1544	2.3059	2.5003	2.7955	3.3682	4.0331
4	2.4472	2.5971	2.7892	3.0802	3.6437	4.2973
5	2.6999	2.8487	3.0391	3.3272	3.8841	4.5293
6	2.9254	3.0735	3.2626	3.5485	4.1002	4.7390
7	3.1310	3.2784	3.4666	3.7506	4.2983	4.9317
8	3.3212	3.4680	3.6553	3.9379	4.4822	5.1112
9	3.4989	3.6453	3.8319	4.1133	4.6547	5.2799
10	3.6663	3.8123	3.9984	4.2787	4.8176	5.4395

doi:10.1371/journal.pone.0118537.t002

Wang B, Shi W, Miao Z (2015) Confidence Analysis of Standard Deviational Ellipse and Its Extension into Higher Dimensional Euclidean Space. PLOS ONE 10(3): e0118537. https://doi.org/10.1371/journal.pone.0118537
https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0118537

IDC HERZLIYA

Wang et al 2015 –
Std ellipses in higher dimension



Wang B, Shi W, Miao Z (2015) Confidence Analysis of Standard Deviational Ellipse and Its Extension into Higher Dimensional Euclidean Space. PLOS ONE 10(3): e0118537. https://doi.org/10.1371/journal.pone.0118537
https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0118537

IDC HERZLIYA

Summary of Gauss related topics

- The normal distribution
- Central limit theorem – characterizing the distribution of averages.
- Approximately normal natural phenomena
- Multi D Gaussians and its defining characteristics and quantities.
- Log Normal distribution – next time

IDC HERZLIYA