MATH F113 (Probability and Statistics)

Chandra Shekhar
Associate Professor

Department of Mathematics
BITS Pilani, Pilani Campus, Rajasthan 333 031
Email: chandrashekhar@pilani.bits-pilani.ac.in
Mobile: 9414492349

In Lecture 17

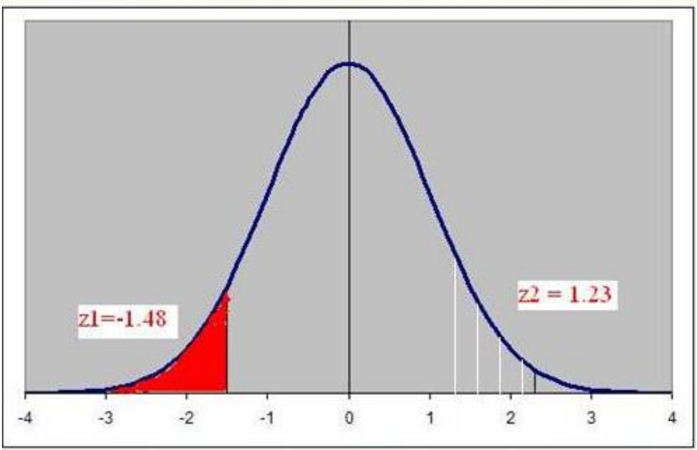
Nomal Distribution
Standard Normal Distribution

Exercise: X **is normally distributed and the mean of** X **is 12 and standard deviation is 4.**

- **Find** $P(X \geq 20), P(X \leq 20), P(0 \leq X \leq 12)$
- **Find** x' **where** $P(X > x') = 0.24$
- **Find** x_0 **and** x_1**, where** $P(x_0 < X < x_1) = 0.50$ **and** $P(X > x_1) = 0.25$

Exercise: In a distribution exactly normal, 7% of the items are under 35 and 89% are under 63. What are the mean and standard deviation of the distribution?

Normal Distribution (Cont...)



Solution:

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

$$P(X < 35) = 0.07, P(X < 63) = 0.89$$

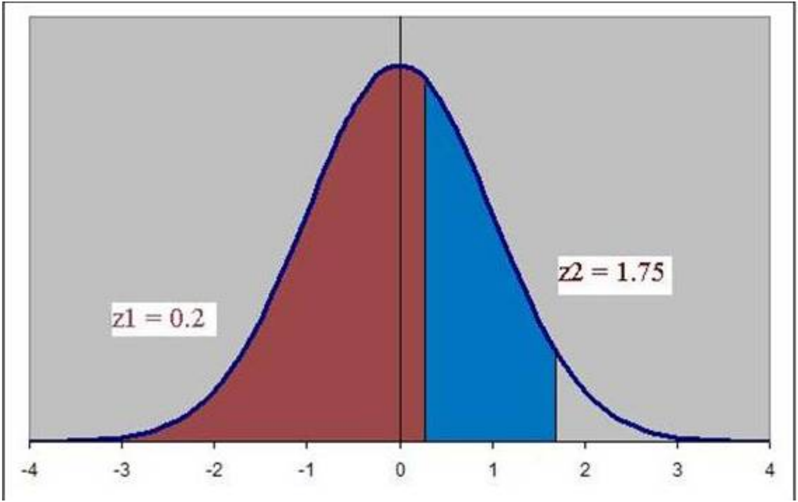
$$P(Z < z_1) = 0.07 \,\textbf{where}\, z_1 = \frac{35 - \mu}{\sigma}$$

$$P(Z < z_2) = 0.89 \,\textbf{where}\, z_2 = \frac{63 - \mu}{\sigma}$$

$$\textbf{From table,}\, z_1 = -1.48, z_2 = 1.23.$$

$$\mu = 50.3, \sigma = 10.33$$

Exercise: If the skulls are classified as A, B, C according as length-breadth index is under 75, between 75 and 80, or over 80, find approximately the mean and standard deviation of a series in which A are 58%, B are 38% and C are 4%.


z1 = 0.2
z2 = 1.75
-4
-3
-2
-1
0
1
2
3
4

Solution:

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

$$P(X < 75) = 0.58, P(X > 80) = 0.04$$

$$P(Z < z_1) = 0.58 \, \textbf{where} \, z_1 = \frac{75 - \mu}{\sigma} ... etc$$

$$z_1 = 0.2, \; z_2 = 1.75$$

Exercise 43/pp.146: Let X denote the time in hours needed to locate and correct a problem in the software that governs the timing of traffic lights in the downtown area of a large city. Assume X is normally distributed with mean 10 hours and variance 9.

(a) Find the probability that the next problem will require at most 15 hours to find and correct.
(b) The fastest 5% of repairs take at most how many hours to complete?

Exercise 41/4.4/pp.146:

Most galaxies take the form of a flattened disc, with the major part of the light coming from this very thin fundamental plane. The degree of flattening differs from galaxy to galaxy. In the milky way Galaxy most gases are concentrated near the center of the fundamental plane.

Let X denote the perpendicular distance from this center to a gaseous mass. Let X is normally distributed with mean 0 and standard deviation 100 parsecs

(a) Sketch a graph of the density for X. Indicate on this graph the probability that a gaseous mass is located within 200 parsecs of the center of the fundamental plane. Find this probability

(b) Approximately, what percentage of the gaseous masses are located more than 250 parsecs from the center of the plane?
(c) What distance has the property that 20% of the gaseous masses are at least this far from the fundamental plane?

Exercise: **Let** $X \sim N(\mu, \sigma^2)$ **if** $\sigma^2 = \mu^2, (\mu > 0)$ **express**

$$P\left(X < -\mu | X < \mu\right)$$

in terms of cumulative distribution function of $N(0, 1)$

Log Normal Distribution **The positive random variable X is said to have a log normal distribution, if $\log_e X$ is normally distributed. That is,**

$$\log_e X \sim N(\mu, \sigma)$$

In probability and statistics, the log-normal distribution is the single tailed probability distribution of any random variable whose logarithm is normally distributed.
If Y is a random variable with a normal distribution, then $X = exp(Y)$ has a log normal distribution; likewise, if X is log normally distributed, then $log(X)$ is normally distributed.

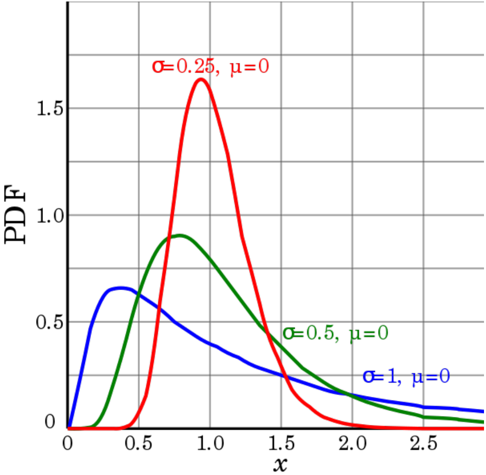
Log Normal Distribution: **The PDF is**

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}; \qquad x > 0$$

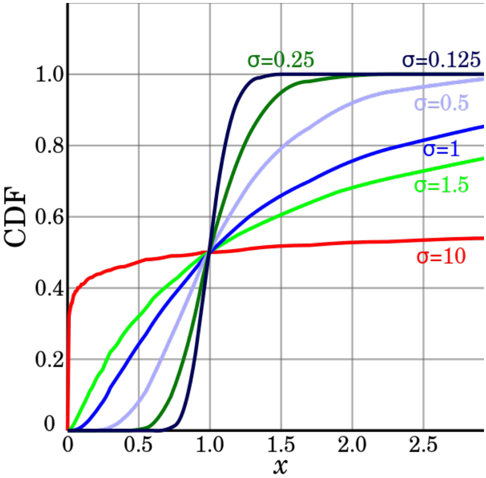
The CDF is

$$F(x) = \frac{1}{2} + \frac{1}{2} erf\left[\frac{\ln x - \mu}{\sigma\sqrt{2}}\right]$$

Log Normal Distribution (Cont...)



Log Normal Distribution (Cont...)



Remark: The moment generating function for log normal does not exist.

The mean is

$$E(X) = e^{\mu + \sigma^2/2}$$

and the variance is

$$Var(X) = \left(e^{\sigma^2} - 1\right) e^{2\mu + \sigma^2}$$

Hence,

$$\mu = \ln(E(X)) - \frac{1}{2}\ln\left(1 + \frac{Var(X)}{(E(X))^2}\right)$$

$$\sigma^2 = \ln\left(\frac{Var(X)}{(E(X))^2} + 1\right)$$

Exercise 45/pp.146 Let X be normal with mean μ and variance σ^2. Let G denote the cumulative distribution for $Y = e^X$ and let F denote the cumulative distribution for X. Then,

$$G(y) = F(\ln y)$$

and

$$G'(y) = F'(\ln y)/y$$

Log Normal Distribution (Cont...)

$$G(y) = P(Y \leq y) = P(e^X \leq y)$$
$$= P(X \leq \ln y)$$

Now $X \sim N(\mu, \sigma)$. **Therefore,**

$$G(y) = P\left(X \leq \ln y\right)$$

$$\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\ln y} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \qquad \mu, x \in (-\infty, \infty)\, \sigma > 0$$

or,

$$G(y) = \frac{1}{\sqrt{2\pi}\sigma} \int_0^y e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}} d(\ln y)$$

$$\mu, x \in (-\infty, \infty) \qquad y, \sigma > 0$$

where $y = e^x$

$$G(y) = \frac{1}{\sqrt{2\pi}\sigma} \int_0^y e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}} d(\ln y) = F(\ln y)$$

Hence show that the density for Y is given by

$$g(y) = \frac{1}{\sqrt{2\pi}\sigma y} e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}},$$

$$\mu \in (-\infty, \infty), y, \sigma > 0$$

$$G(y) = \frac{1}{\sqrt{2\pi}\sigma} \int_0^y e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}} d(\ln y) = F(\ln y)$$

$$g(y) = G'(y) = F'(\ln y)/y$$

Exercise 46/pp.147 **Let Y denote the diameter in millimeters of styrofoam pellets used in packing. Assume that Y has a log-normal distribution with parameter $\mu = 0.8$, $\sigma = 0.1$.**

(i) Find the probability that a randomly selected pellet has a diameter that exceeds 2.7 mm

(ii) Between what two values will Y fall with probability approximately 0.95?

Solution: (i)

$$P(Y > 2.7) = P(Z > (\ln 2.7 - 0.8)/0.1)$$
$$= 1 - F(1.93)$$

(ii) We need to find y_1 and y_2 such that

$$P(y_1 < Y < y_2) = 0.95$$

Find symmetric to mean

$$P\left(\frac{\ln y_1 - 0.8}{0.1}) < Z < \frac{\ln y_2 - 0.8}{0.1}\right) = 0.95$$

$$(\ln y_1 - 0.8)/0.1 = z_{0.975} = -1.96, y_1 = 1.829$$

$$(\ln y_2 - 0.8)/0.1 = z_{0.025} = 1.96, y_2 = 2.707$$

Exercise: **If $\log_{10} X$ is normally distributed with mean 4 and variance 4, find**

$$P(1.202 < X < 83180000)$$

$$(Given, \log_{10} 1202 = 3.08, \log_{10} 8318 = 3.92)$$

Solution: **Since,** $\log X$ **is an increasing function of** X**, we have**

$$P(1.202 < X < 83180000)$$

$$= P(\log_{10} 1.202 < \log_{10} X < \log_{10} 83180000)$$

$$= P(0.08 < \log_{10} X < 7.92)$$

$$= P(0.08 < Y < 7.92)$$

where

$$Y = \log_{10} X \sim N(4, 2)$$

Required Probability

$$= P(0.08 < Y < 7.92)$$
$$= P(-1.96 < Z < 1.96)$$
$$= 2P(0 < Z < 1.96) = 0.95$$