Normal Probability Rules

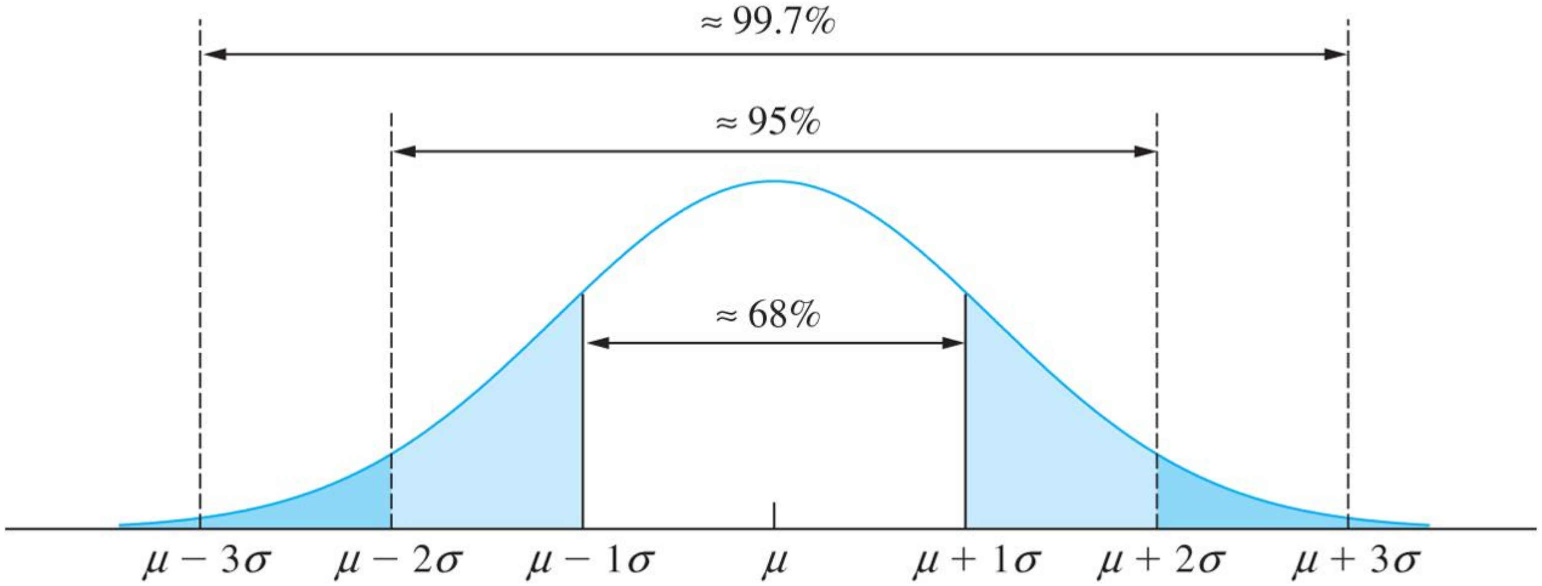
Let X be a normal random variable with parameters $\mu\ and\ \sigma$. Then

$$P[-\sigma < X - \mu < \sigma] \approx 0.68$$
$$P[-2\sigma < X - \mu < 2\sigma] \approx 0.95$$
$$P[-3\sigma < X - \mu < 3\sigma] \approx 0.997$$

$$P[-\sigma < X - \mu < \sigma] = P\left[-1 < \frac{X-\mu}{\sigma} < 1\right] = P[-1 < Z < 1] \approx 0.68$$
$$P[-2\sigma < X - \mu < 2\sigma] = P\left[-2 < \frac{X-\mu}{\sigma} < 2\right] = P[-2 < Z < 2] \approx 0.95$$
$$P[-3\sigma < X - \mu < 3\sigma] = P\left[-3 < \frac{X-\mu}{\sigma} < 3\right] = P[-3 < Z < 3] \approx 0.997$$

If X is a normal variable, we can be 99.7% sure that a value x will be within 3σ distance from the center μ

It is 0.3% rare that a value of a normal variable can be 3σ away from the center μ.



Chebyshev's Inequality

Let X be a random variable with mean μ and standard deviation σ. Then for any positive number k,

$$P[|X-\mu| < k\sigma] \geq 1 - \frac{1}{k^2}$$

$$k = 1:\ \ P[|X-\mu| < \sigma] = P[-\sigma < X-\mu < \sigma] \geq 1 - 1 = 0$$
$$k = 2:\ \ P[|X-\mu| < 2\sigma] = P[-2\sigma < X-\mu < 2\sigma] \geq 1 - \tfrac{1}{4} = 0.75$$
$$k = 3:\ \ P[|X-\mu| < 3\sigma] = P[-3\sigma < X-\mu < 3\sigma] \geq 1 - \tfrac{1}{9} = 0.89$$

This result is for any random variable.

Example: Check the normal probability rules using the Chebyshev's Inequality

$$P[|X-\mu| < \sigma] = P[-\sigma < X-\mu < \sigma] \geq 1 - 1 = 0 \ \ vs\ \ 0.68$$
$$P[|X-\mu| < 2\sigma] = P[-2\sigma < X-\mu < 2\sigma] \geq 1 - \frac{1}{4} = 0.75 \ \ vs\ \ 0.95$$
$$P[|X-\mu| < 3\sigma] = P[-3\sigma < X-\mu < 3\sigma] \geq 1 - \frac{1}{9} = 0.89 \ \ vs\ \ 0.997$$

The result from Chebyshev's inequality is more conservative.

Chebyshev’s Inequality

Example: Let M denote the total staffing-hours working without a serious accident. Past experience indicates that M has a mean of 2 million and a standard deviation of 0.1 million. A serious accident has just occurred. Would it be unusual for the next serious accident to occur within the next 1.6 million staffing-hours?

Solution:
We need to find the probability $P[M \leq 1.6]$ with the distribution of M unknown!!!
However, from Chebyshev’s inequality with $k = 4$, we have,

$$P[|M-2| < 0.4] \geq 1 - \frac{1}{16} = 0.9375$$
$$\Rightarrow P[1.6 < M < 2.4] \geq 0.9375$$

This implies that: $P[M \leq 1.6] + P[M \geq 2.4] \leq 0.0625$
Since $P[M \geq 2.4] \geq 0$, we have,$P[M \leq 1.6] \leq 0.0625$
Therefore, it is unusual.

If it is known that the density of M is symmetric, can this probability be improved?

Since the density if symmetric, we have$P[M \leq 1.6] = P[M \geq 2.4]$. Therefore, $P[M \leq 1.6] \leq \frac{0.0625}{2} = 0.03125$

Continuous Distributions

Distributions	Density function	Moment generating function	Mean	Variance
Gamma	$f(x)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{\alpha-1}e^{-\frac{x}{\beta}},$ $x>0,\alpha>0,\beta>0$	$(1-\beta t)^{-\alpha}$	$\alpha\beta$	$\alpha\beta^2$
Exponential	$f(x)=\frac{1}{\beta}e^{-\frac{x}{\beta}},\qquad x>0,\beta>0$	$(1-\beta t)^{-1}$	β	β^2
Chi-Squared	$f(x,\gamma)=\frac{x^{(\frac{\gamma}{2}-1)}e^{-x/2}}{2^{\frac{\gamma}{2}}\Gamma(\frac{\gamma}{2})},x>0$	$(1-2t)^{-\gamma/2}$	γ	2γ
Normal	$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$	$e^{(\mu t+\frac{1}{2}\sigma^2t^2)}$	μ	σ^2

Joint Distribution

Joint density of discrete random variables:

Let X and Y be discrete random variables. The ordered pair (X, Y) is called a two dimensional discrete random variable.

A function $f_{XY}(\mathrm{x}, \mathrm{y}) = P[X = x \; and \; Y = y]$ is called the joint density for (X, Y) if

$$f_{XY}(x, y) \geq 0 \; and \; \sum_{all\, x} \sum_{all\, y} f_{XY}(x, y) = 1$$

Example:

In an automobile plant, one robot is welding two joints. The second robot is tightening three bolts. Let X denote the number of defective welds and Y the number of improperly tightened bolts produced per car. The joint density function for (X, Y) is shown in the table.

X\Y	**0**	**1**	**2**	**3**
0	0.840	0.030	0.020	0.010
1	0.060	0.010	0.008	0.002
2	0.010	0.005	0.004	0.001

1. Verify that the joint density function in the table is a discrete joint density function
2. Calculate the probability that there will be exactly one error made
3. Calculate the probability that there will be no improperly tightened bolts

Joint Distribution

Marginal densities of discrete random variables:

Let (X, Y) be a two dimensional discrete random variable with joint density $f_{XY}(\mathrm{x}, \mathrm{y})$.
The marginal density for X, denoted by f_X, is given by $f_X(x) = \sum_{all\ y} f_{XY}(x, y)$
The marginal density for Y, denoted by f_Y, is given by $f_Y(y) = \sum_{all\ x} f_{XY}(x, y)$

Example:

In an automobile plant, one robot is welding two joints. The second robot is tightening three bolts. Let X denote the number of defective welds and Y the number of improperly tightened bolts produced per car. The joint density function for (X, Y) is shown in the table. Find the marginal density f_X and f_Y.

X\Y	**0**	**1**	**2**	**3**
0	0.840	0.030	0.020	0.010
1	0.060	0.010	0.008	0.002
2	0.010	0.005	0.004	0.001

How about (X, Y, Z)?

Joint Distribution

Joint density of continuous random variables:
Let X and Y be continuous random variables. The ordered pair (X, Y) is called a two dimensional continuous random variable.
A function $f_{XY}(\mathrm{x}, \mathrm{y})$ is called the joint density for (X, Y) if $f_{XY}(x, y) \geq 0.\, and\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)dydx = 1$

$$P[a \leq X \leq b\ and\ c \leq Y \leq d] = \int_{a}^{b}\int_{c}^{d} f_{XY}(x, y)dydx$$

Where , a, b, c, d are real constants.

Practice Example: The joint density for (X, Y) is $f_{XY}(x, y) = c,\ \ 8.5 \leq x \leq 10.5, 120 \leq y \leq 240.$

(1) Find the value of c so that $f_{XY}(x, y)$ is a valid density.
(2) Calculate $P[9 \leq x \leq 10\ and\ 125 \leq y \leq 140]$

(1) We want $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)dydx = 1 \Rightarrow \int_{8.5}^{10.5}\int_{120}^{240} cdydx = 1 \Rightarrow 240c = 1 \Rightarrow c = \frac{1}{240}$

(2) $P[9 \leq x \leq 10\ and\ 125 \leq y \leq 140] = \int_{9.0}^{10.0}\int_{125}^{140} \frac{1}{240} dydx = \frac{1}{16}$

Joint Distribution

Marginal densities of continuous random variables:

Let (X, Y) be a two dimensional continuous random variable with joint density $f_{XY}(\mathrm{x}, \mathrm{y})$.

The *marginal density* for X, denoted by f_X, is given by $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy$

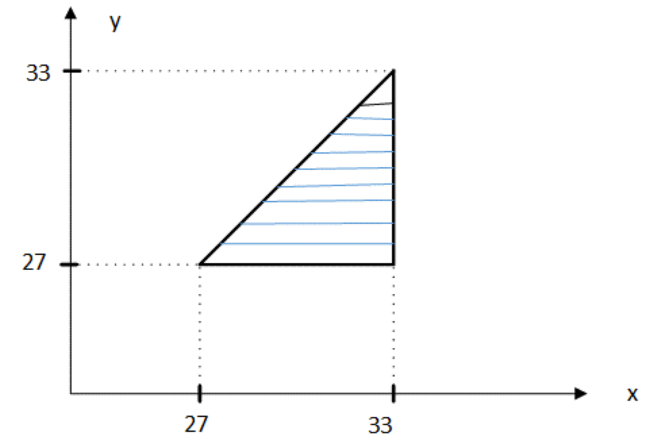
The *marginal density* for Y, denoted by f_Y, is given by $f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)dx$

Example:

Assume that the joint density for (X, Y) is given by $f_{XY}(x, y) = \frac{c}{x}, 27 \le y \le x \le 33$

(1) Find the value of c so that $f_{XY}(x, y)$ is a valid density function

(2) Find the marginal density $f_X(x)\ and\, f_Y(y)$



(1) In order for $f_{XY}(x, y)$ to be a valid density function, we need,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)dydx = 1 \Rightarrow \int_{27}^{33}\int_{27}^{x}\frac{c}{x}dydx = 1 \Rightarrow \int_{27}^{33}\frac{c}{x}(x-27)dx = 1$$

$$\Rightarrow c\int_{27}^{33}(1-\frac{27}{x})dx = 1 \Rightarrow c\left[6 - 27\ln\left(\frac{33}{27}\right)\right] = 1 \Rightarrow c = \frac{1}{6-27\ln\left(\frac{33}{27}\right)} \approx 1.72$$

(2) $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy = \int_{27}^{x}\frac{c}{x}dy = c\left(1-\frac{27}{x}\right), 27 \le x \le 33$

$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)dx = \int_{y}^{33}\frac{c}{x}dx = c(ln33 - lny),\ 27 \le y \le 33$