Chapter 1. Probability

Microscopic properties of matter: quantum mechanics, atomic and molecular properties

Macroscopic properties of matter: thermodynamics, E, H, C_V, C_p, S, A, G

How do we relate these two properties?

Statistical thermodynamics (mechanics)

Basic Probability Theory

Variables: quantities that can change in value throughout the course of an experiment or series of events
e.g., the side of coin observed after tossing the coin.

Discrete variables: – assume only a limited number of specific values
e.g., the outcome of toss = two values (head or tail)
– sample space of the variables = the possible values a variable can assume
e.g., the outcome of toss {+1, –1}

Continuous variables:
– assume any value in the certain range
e.g., temperature, $0 < T < \infty$.

Imagine a lottery where balls numbered 1 to 50 are randomly mixed.

– The probability of selecting ① is 1/50.
– This requires an infinite number of experiments.

Consider a variable *x* for which the sample space consists of *n* values denoted as $\{x_1, x_2, \cdots x_n\}$.

– The probability that a variable *x* will assume one of these values (P_i) is: $0 < P_i < 1$, $i = 1, 2, \cdots, M$

– The sum of the probabilities for selecting each individual ball must be equal to 1.

$$P_1 + P_2 + \cdots + P_n = \sum_{i=1}^{M} P_i = 1$$

Consider the probability associated with a given outcome for a series of experiments, i.e, the event probability.

- Imagine tossing a coin four times.
- What is the probability that at least two heads are observed after four tosses?

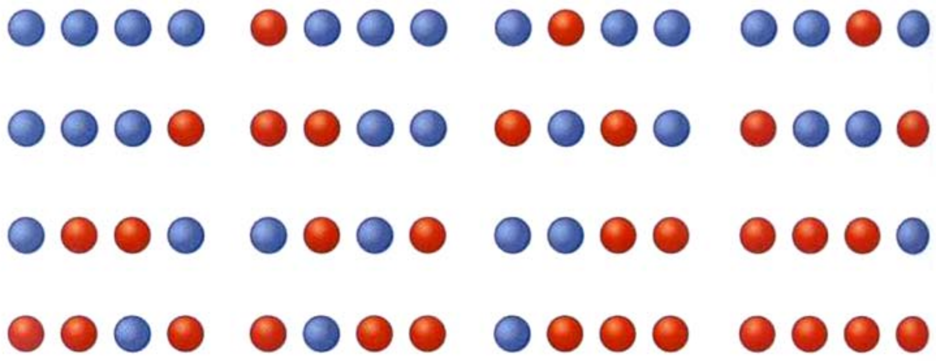


Figure 1 Potential outcomes after tossing a coin four times. Red signifies heads and blue signifies tails.

- The probability = 11/16
- The probability (P_E) that the outcome or event of interest, E, occurs in N values in sample space

$$P_E = \frac{E}{N}$$

The Fundamental Counting Principle

For a series of manipulations $\{M_1, M_2, \cdots, M_j\}$ having n_j ways to accomplish each manipulation $\{n_1, n_2, \cdots, n_j\}$, the total number of ways to perform the entire series of manipulations (Total_M) is the product of the number of ways to perform each manipulation under the assumption that the ways are independent:

$$\text{Total}_M = (n_1)(n_2)(n_3)\cdots(n_j)$$

- Assemble 30 students in a line.
- How many arrangements of students are possible?

The total number of ways

$$W = 30 \times 29 \times 28 \times \cdots \times 2 \times 1 = 30! = 2.65 \times 10^{32}$$

Example:

How many five-card arrangements are possible from a standard deck of 52 cards?

Solution:

$\text{Total}_M = (n_1)\ (n_2)\ (n_3)\ (n_4)\ (n_5) = (52)\ (51)\ (50)\ (49)\ (48) = 311{,}875{,}200$

Permutations

– How many permutations are possible if only a subset of objects is employed in constructing the permutation?

– $P(n, j)$: the number of permutations possible using a subset of j objects from the total group of n

$$P(n, j) = n(n-1)(n-2)\cdots(n-j+1)$$
$$= \frac{n(n-1)(n-2)\cdots(2)(1)}{(n-j)(n-j-1)\cdots(2)(1)} = \frac{n!}{(n-j)!}$$

Example:

The coach of a basketball team has 12 plays on the roster, but can only play 5 plays on one time. How many 5-player arrangements are possible using the 12-player roster?

Solution:

$$P(n, j) = P(12, 5) = \frac{12!}{(12-5)!} = 95,040$$

Configurations

– permutations = the number of ordered arrangements
– configurations = the number of unordered arrangements

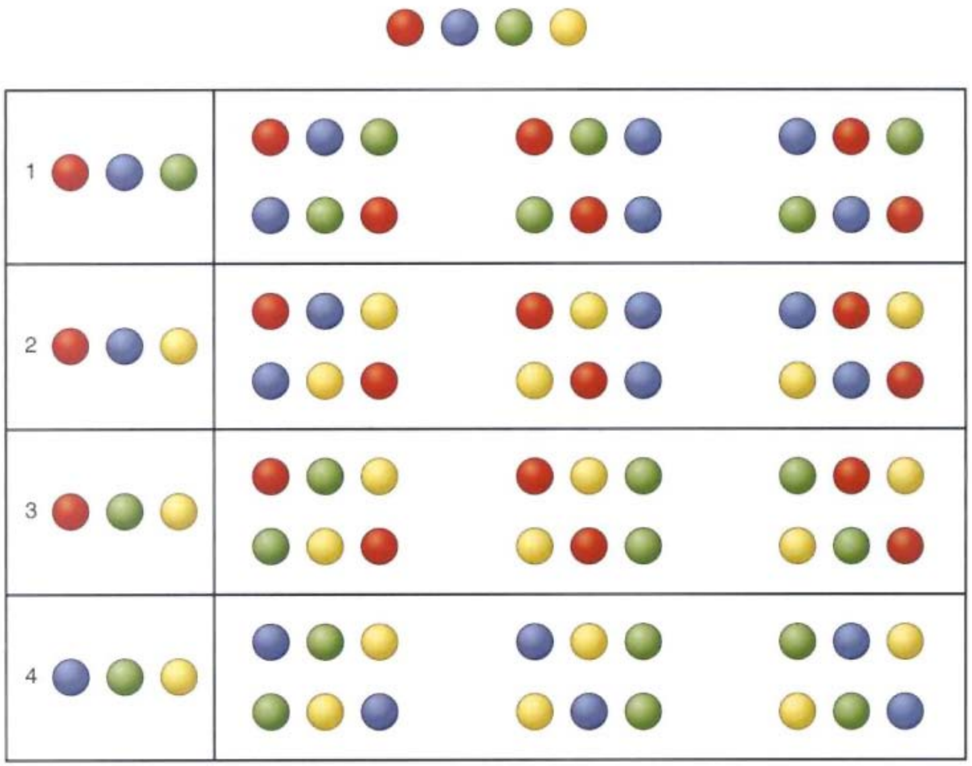


Figure 2 Illustration of configurations and permutations using four colored balls. The left-hand column presents the four possible three-color configurations, and the right-hand column presents the six permutations corresponding to each configuration.

$C(n, j)$ = the number of configurations that are possible using a subset of j objects from a total number of n objects.

$$C(n,j) = \frac{P(n,j)}{j!} = \frac{n!}{j!(n-j)!}$$

Example:

How many possible 5–card combinations or “ hands” from a standard 52–card deck are there?

Solution:

$$C(52,5) = \frac{52!}{5!(52-5)!} = 2,598,960$$

Binomial Probabilities

– Define the complement of P_E as the probability of an outcome other than that associated with the event of interest, as denoted by P_{EC}

$$\therefore P_E + P_{EC} = 1$$

– Bernoulli trial: the outcome of a given experiment will be a success (i.e., the outcome of interest) or a failure (i.e., not the outcome of interest).

– Binomial experiment: a collection of Bernoulli trials

– The probability of observing heads every time when a coin is tossed four times

$$P_E = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \left(\frac{1}{16}\right)$$

– The probability of obtaining j successes in a trial consisting of n trials for a series of Bernoulli trials in which the probability of success for a single trial is P_E:

$$P(j) = C(n, j)(P_E)^j (1 - P_E)^{n-j}$$

$$= \frac{n!}{j!(n-j)!}(P_E)^j (1 - P_E)^{n-j}$$

$C(n, j)$: the number of configurations that are possible using a subset of j successes in n trials

Example:

Toss a coin 50 times. What are the probabilities of having the coin land heads up 10 times and 25 times?

Solution:

$$P_{10} = C(n, j)(P_E)^j (1 - P_E)^{n-j}$$

$$= C(50,10)\left(\frac{1}{2}\right)^{10}\left(\frac{1}{2}\right)^{40}$$

$$= \frac{50!}{(10!)(40!)}\left(\frac{1}{2}\right)^{10}\left(\frac{1}{2}\right)^{40} = 9.1 \times 10^{-6}$$

$$P_{25} = C(50,25)\left(\frac{1}{2}\right)^{25}\left(\frac{1}{2}\right)^{25}$$

$$= \frac{50!}{(25!)(25!)}\left(\frac{1}{2}\right)^{25}\left(\frac{1}{2}\right)^{25} = 0.11$$

Stirling's Approximation

– Calculations of factorial quantities becomes extremely large:

$$100! = 9.3\times 10^{157}$$

– Need an approximation: Stirling's approximation

$$\mathbf{\ln N! = N \ln N - N}$$

– Derivation

$$\begin{aligned}\ln N! &= \ln[(N)(N-1)(N-2)\cdots(2)(1)] \\ &= \ln N + \ln(N-1) + \ln(N-2) + \cdots + \ln 2 + \ln 1 \\ &= \sum_{n=1}^{N} \ln n \approx \int_1^N \ln n \, dn \\ &= N \ln N - N - (1\ln 1 - 1) \approx N \ln N - N\end{aligned}$$

Probability Distribution Functions

Number of Heads	Probability	Number of Heads	Probability
0	8.88×10^{-16}	30	0.042
1	4.44×10^{-14}	35	2.00×10^{-3}
2	1.09×10^{-12}	40	9.12×10^{-6}
5	1.88×10^{-9}	45	1.88×10^{-9}
10	9.12×10^{-6}	48	1.09×10^{-12}
15	2.00×10^{-3}	49	4.44×10^{-14}
20	0.042	50	8.88×10^{-16}
25	0.112		

– This information can be presented graphically by plotting the probability as a function of outcome.

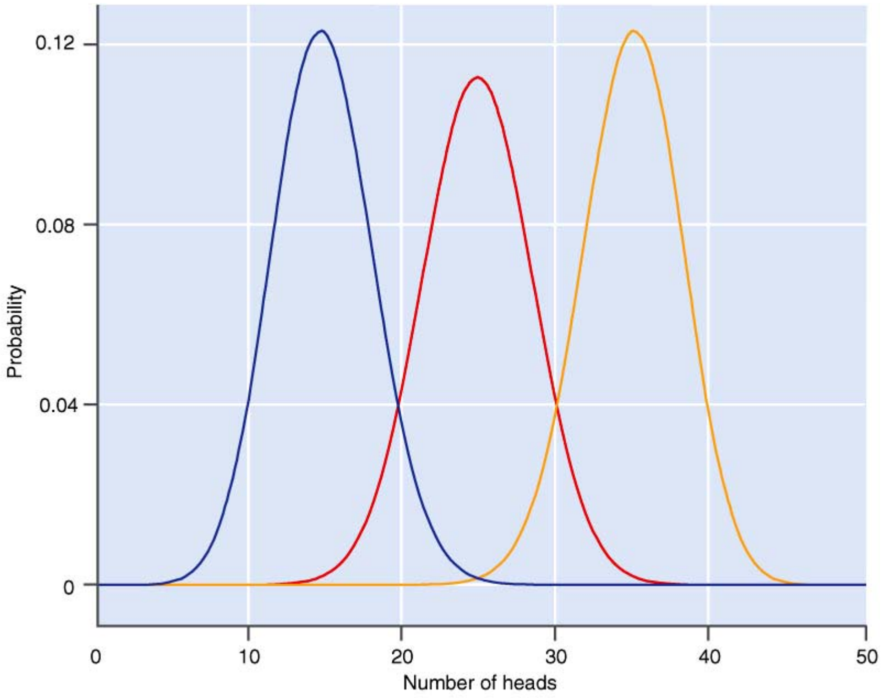


Figure 3. Plot of the probability of the number of heads being observed after flipping a coin 50 times. The red curve represents the distribution of probabilities for P_E=0.5, the blue curve for P_E=0.3, and the yellow curve for P_E=0.7

– The probability of observing *j* successful trials following *n* total trials

$$P(j) = \frac{n!}{j!(n-j)!}(P_E)^j(1-P_E)^{n-j} \qquad j = 0, 1, 2, \cdots, n$$

– A probability distribution function (f) represents the probability of a variable (x) having a given value, with the probability described by a function $P(x_i) \propto f_i$.

$$P(x_i) = Cf_i$$

$$\sum_{i=1}^{M} P(x_i) = 1$$

$$1 = \sum_{i=1}^{M} Cf_i = Cf_1 + Cf_2 + \cdots + Cf_M = C\sum_{i=1}^{M} f_i$$

$$C = \frac{1}{\sum_{i=1}^{M} f_i}$$

$$\therefore P(x_i) = \frac{f_i}{\sum_{i=1}^{M} f_i}$$

Probability Distributions Involving Discrete and Continuous Variables

– If the variable x is continuous, $P(x)$ is the probability that the variable x has a value in the range of dx

$P(x)dx = Cf(x)dx$, where $f(x)$ is a function not yet defined.

$$\int_{x_1}^{x_2} P(x)dx = C\int_{x_1}^{x_2} f(x)dx = 1$$

$$C = \frac{1}{\int_{x_1}^{x_2} f(x)dx}$$

$$P(x)dx = Cf(x)dx = \frac{f(x)dx}{\int_{x_1}^{x_2} f(x)dx}$$

Characterizing Distribution Functions

Average Values

– Consider a function, $g(x)$, whose value is dependent on x.

$$\langle g(x)\rangle = \sum_{i=1}^{M} g(x_i) P(x_i) = \frac{\sum_{i=1}^{M} g(x_i)\, f_i}{\sum_{i=1}^{M} f_i}$$

Distribution moments

– $g(x) \propto x^n$

– $\langle x \rangle$: the first moment of the distribution function

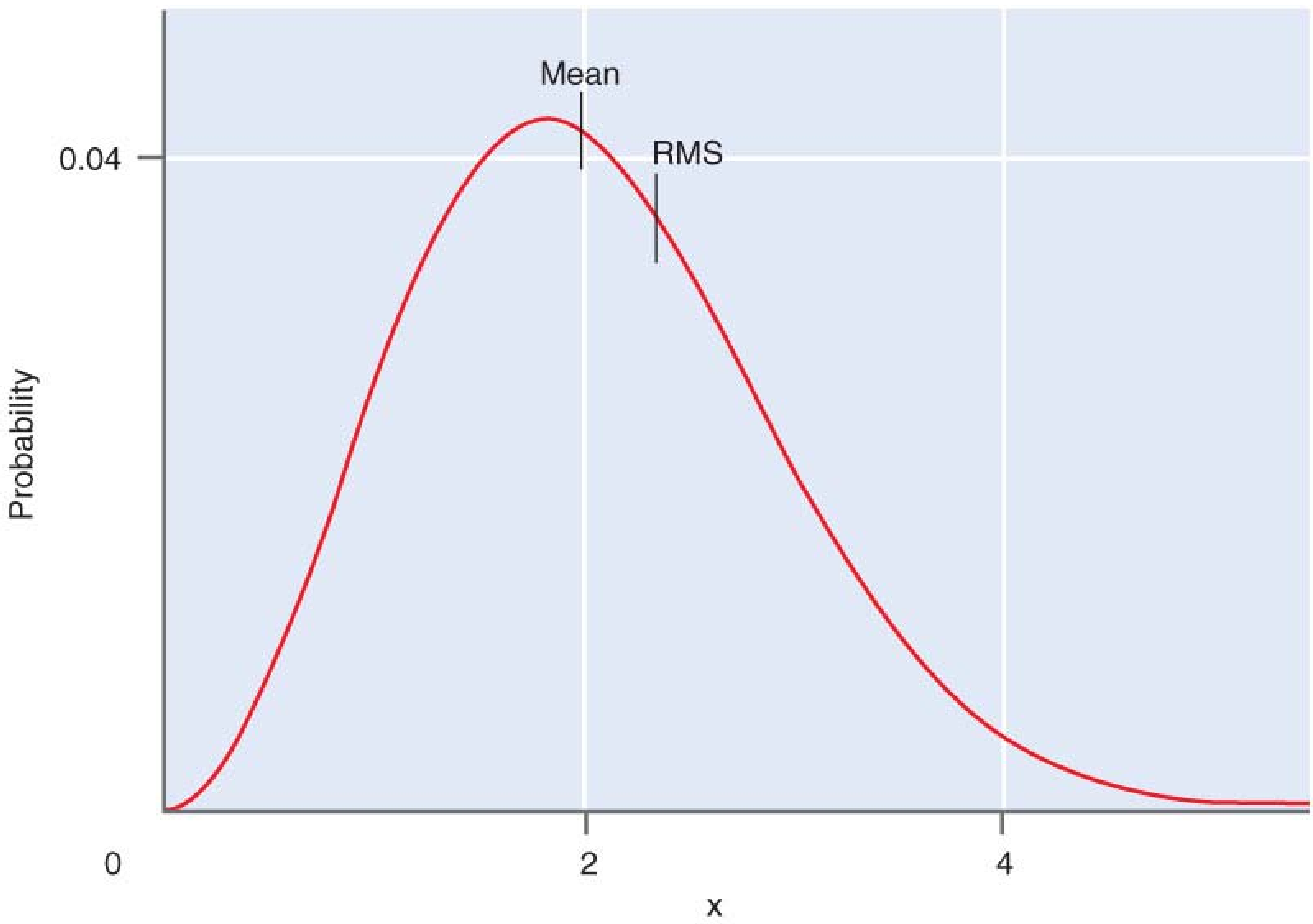
– $\langle x^2 \rangle$: the second moment of the distribution

– $\sqrt{\langle x^2 \rangle}$: the root-mean-squared (*rms*) value

Example

$$P(x) = Cx^2 e^{-ax^2} \qquad 0 \le x \le \infty$$

Are the mean and *rms* values for this distribution the same?


Mean
RMS
0.04
Probability
0
2
4
x

Variance

– The variance (σ^2) : a measure of the width of a distribution defined as the average deviation squared from the mean of distribution

$$\sigma^2 = \left\langle \left(x - \langle x \rangle \right)^2 \right\rangle = \left\langle x^2 - 2x\langle x \rangle + \langle x \rangle^2 \right\rangle$$

Note:

$$\langle b(x) + d(x) \rangle = \langle b(x) \rangle + \langle d(x) \rangle$$

$$\langle cb(x) \rangle = c\langle b(x) \rangle$$

$$\therefore \sigma^2 = \left\langle x^2 - 2x\langle x \rangle + \langle x \rangle^2 \right\rangle = \langle x^2 \rangle - 2\langle x \rangle \langle x \rangle + \langle x \rangle^2$$

$$= \langle x^2 \rangle - \langle x \rangle^2$$

- Gaussian distribution: the "bell-shaped curve"

$$P(x)dx = \frac{1}{\sqrt{(2\pi\sigma^2)}} e^{-\frac{(x-\delta)^2}{2\sigma^2}} dx$$

σ^2: the width of the distribution

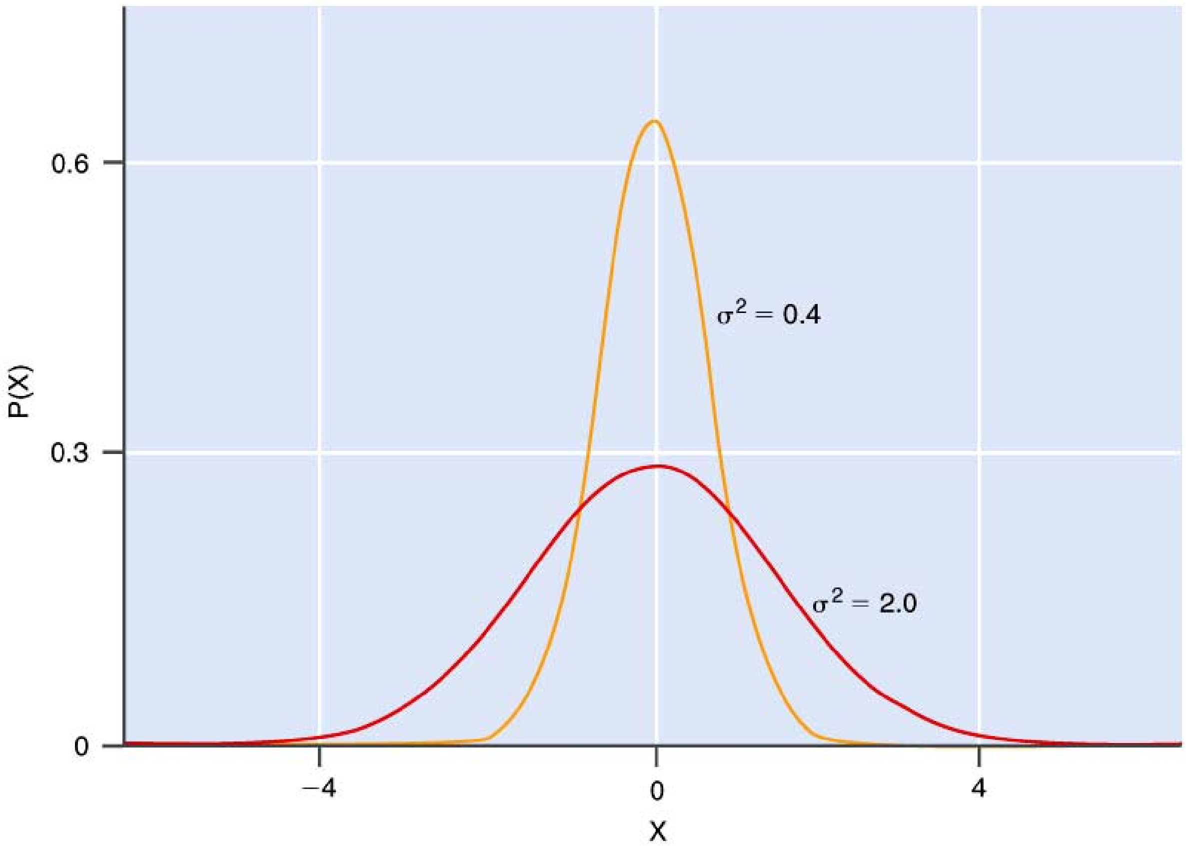


Figure 4. The influence of variance on Gaussian probability distribution functions. Notice that an increase in the variance corresponds to an increase in the width of the distribution.

Example:

$$P(x) = Cx^2 e^{-ax^2} \qquad \sigma^2 = ?$$