Lecture 04.08 Probability density and mass functions

frequency distribution
probability mass function

Consider an experiment that measures a random variable. We can plot the relative frequency of the measurand landing in different "bins" (ranges of values). This is called a *frequency distribution* or a *probability mass function* (PMF).

Consider, for instance, a probability mass function as plotted in Figure 04.3, where a frequency a_i can be interpreted as an estimate of the probability of the measurand being in the ith interval. The sum of the frequencies must be unity:

with k being the number of bins.

frequency density distribution

The *frequency density distribution* is similar to the frequency distribution, but with $a_i \mapsto a_i/\Delta x$, where Δx is the bin width.

probability density function

If we let the bin width approach zero, we derive the *probability density function* (PDF)

$$f(x) = \lim_{\substack{k \to \infty \\ \Delta x \to 0}} \sum_{j=1}^{k} a_j/\Delta x. \tag{04.3}$$

We typically think of a probability density function f, like the one in Figure 04.4 as a function that can be integrated over to find the probability of the random variable (measurand) being in an interval $[a, b]$:

$$P(x \in [a, b]) = \int_a^b f(\chi)\,d\chi. \tag{04.4}$$

Of course,

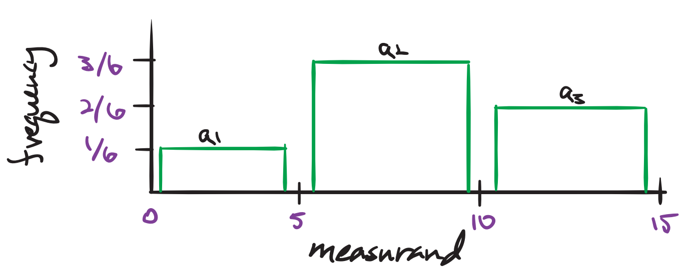

Figure 04.3: plot of a probability mass function.

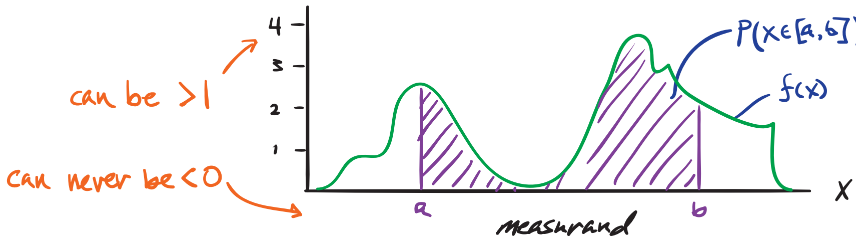

Figure 04.4: plot of a probability density function.

We now consider a common PMF and a common PDF.

04.08.1 Binomial PMF

Consider a random binary sequence of length n such that each element is a random 0 or 1, generated independently, like

$$(1, 0, 1, 1, 0, \cdots, 1, 1). \tag{04.5}$$

Let events $\{1\}$ and $\{0\}$ be mutually exclusive and exhaustive and $P(\{1\}) = p$. The probability of the sequence above occurring is

There are n choose k,

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \tag{04.6}$$

possible combinations of k ones for n bits. Therefore, the probability of any combination of k ones in a series is

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k}. \tag{04.7}$$

We call Equation 04.7 the *binomial distribution PDF*.

binomial distribution PDF

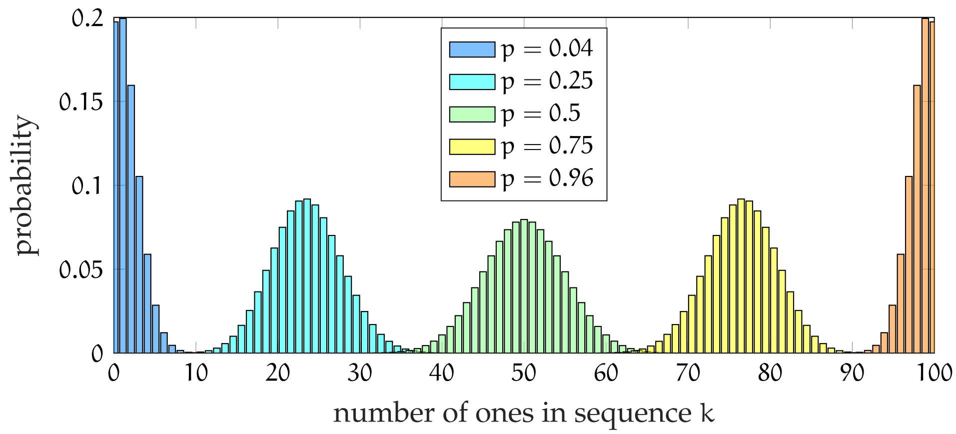

Figure 04.5: binomial PDF for $n = 100$ measurements and different values of $P(\{1\}) = p$, the probability of a measurement error. The plot is generated by the *Matlab* code of Figure 04.6.

Example 04.08-1 Binomial PMF

Consider a field sensor that fails for a given measurement with probability p. Given n measurements, plot the binomial PMF as a function of k failed measurements for a few different probabilities of failure $p \in [0.04, 0.2, 0.5]$.

Figure 04.6 shows *Matlab* code for constructing the PDFs plotted in Figure 04.5. Note that the symmetry is due to the fact that events $\{1\}$ and $\{0\}$ are mutually exclusive and exhaustive.

04.08.2 Gaussian PDF

Gaussian or normal random variable

The *Gaussian* or *normal random variable* x has PDF

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(x-\mu)^2}{2\sigma^2}. \tag{04.8}$$

Although we're not quite ready to understand these quantities in detail, it can be shown that the parameters μ and σ have the following meanings:

mean
standard deviation
variance

- μ is the *mean* of x,
- σ is the *standard deviation* of x, and
- σ^2 is the *variance* of x.

```
%% parameters
n = 100;
k_a = linspace(1,n,n);
p_a = [.04,.25,.5,.75,.96];

%% binomial function
f = @(n,k,p) nchoosek(n,k)*p^k*(1-p)^(n-k);

% loop through to construct an array
f_a = NaN*ones(length(k_a),length(p_a));
for i = 1:length(k_a)
    for j = 1:length(p_a)
        f_a(i,j) = f(n,k_a(i),p_a(j));
    end
end

%% plot
figure
colors = jet(length(p_a));
for j = 1:length(p_a)
    bar(...
        k_a,f_a(:,j),...
        'facecolor',colors(j,:),...
        'facealpha',0.5,...
        'displayname', ['$p = ',num2str(p_a(j)),'$']...
    ); hold on
end
leg = legend('show','location','north');
set(leg,'interpreter','latex')
hold off
xlabel('number of ones in sequence k')
ylabel('probability')
xlim([0,100])
```

Figure 04.6: a *Matlab* script for generating binomial PMFs.

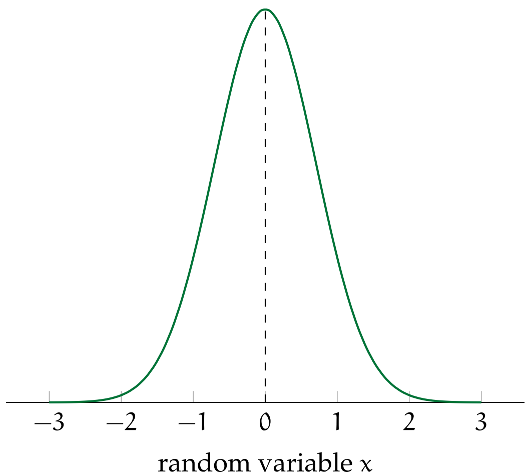

Figure 04.7: PDF for Gaussian random variable x, mean $\mu = 0$, and standard deviation $\sigma = 1/\sqrt{2}$.

Consider the "bell-shaped" Gaussian PDF in Figure 04.7. It is always symmetric. The mean μ is its central value and the standard deviation σ is directly related to its width. We will continue to explore the Gaussian distribution in the following lectures, especially in Lecture 04.12.