prob.pxf Probability density and mass functions

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 Fig. pxf.1, 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.

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.

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

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

We typically think of a probability density function f, like the one in Fig. pxf.2 as a function that can be integrated over to find the probability of the random variable (measurand) being in an interval $[a, b]$:

$$P(x \in [a, b]) = \int_a^b f(\chi)\, d\chi. \tag{2}$$

Of course,

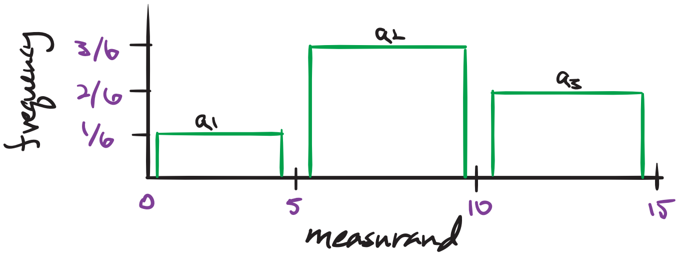

Figure pxf.1: plot of a probability mass function.

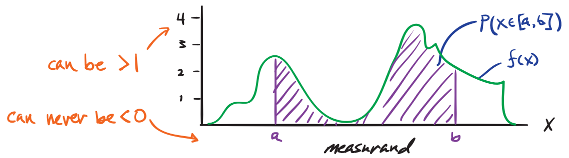

Figure pxf.2: plot of a probability density function.

We now consider a common PMF and a common PDF.

Binomial PMF

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

$$(1, 0, 1, 1, 0, \cdots, 1, 1). \tag{3}$$

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,

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \tag{4}$$

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

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k}. \tag{5}$$

We call Eq. 5 the **binomial distribution PDF**.

Example prob.pxf-1 re: 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 $p \in [0.04, 0.25, 0.5, 0.75, 0.96]$.

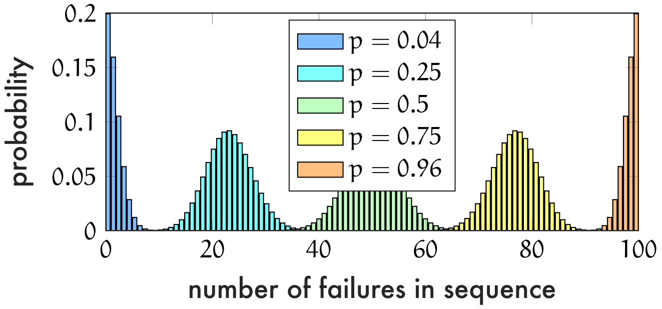

Figure pxf.3: binomial PDF for $n = 100$ measurements and different values of $P(\{1\}) = p$, the probability of a measurement error. The plot is generated by the *Matlab* code of Fig. pxf.4.

Fig. pxf.4 shows *Matlab* code for constructing the PDFs plotted in Fig. pxf.3. Note that the symmetry is due to the fact that events $\{1\}$ and $\{0\}$ are mutually exclusive and exhaustive.

Example prob.pxf-2 **re: hi**

Sed mattis, erat sit amet gravida malesuada, elit augue egestas diam, tempus scelerisque nunc nisl vitae libero. Sed consequat feugiat massa. Nunc porta, eros in eleifend varius, erat leo rutrum dui, non convallis lectus orci ut nibh. Sed lorem massa, nonummy quis, egestas id, condimentum at, nisl. Maecenas at nibh. Aliquam et augue at nunc pellentesque ullamcorper. Duis nisl nibh, laoreet suscipit, convallis ut, rutrum id, enim. Phasellus odio. Nulla nulla elit, molestie non, scelerisque at, vestibulum eu, nulla. Ut odio nisl, facilisis id, mollis et, scelerisque nec, enim. Aenean sem leo, pellentesque sit amet, scelerisque sit amet, vehicula pellentesque, sapien.

```
%% 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 pxf.4: a *Matlab* script for generating binomial PMFs.

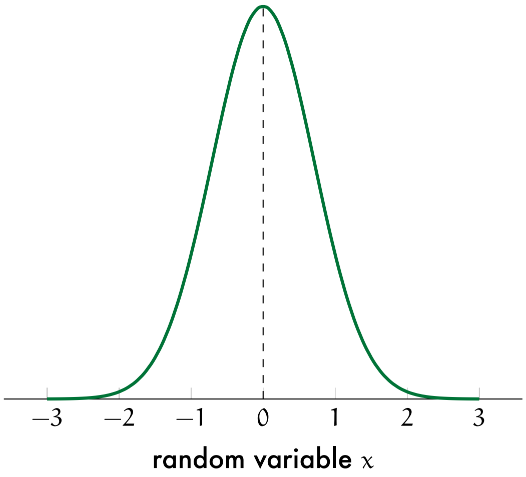

Figure pxf.5: PDF for Gaussian random variable x, mean $\mu = 0$, and standard deviation $\sigma = 1/\sqrt{2}$.

Gaussian PDF

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

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(x-\mu)^2}{2\sigma^2}. \tag{6}$$

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:

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

Consider the "bell-shaped" Gaussian PDF in Fig. pxf.5. 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 Lec. stats.confidence.