This is the matrix we have so far, with the last element missing:

\[
\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i
\]

\(~\text{typical } x\) (based on sample)

\[
S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2
\]

Sample std. dev. = \(S\).

\(~\text{typical deviation/spread}\) (based on sample).

\[
\mu_x = E[x] = \int_{-\infty}^{\infty} x \cdot f(x) \, dx
\]

\(~\text{typical } x\) (based on pop/dist)

\[
\sigma_x^2 = V[x]
\]

\[
\sigma_x^2 = \int_{-\infty}^{\infty} (x - E[x])^2 \cdot f(x) \, dx
\]

Don't drop this \(x\), like the book does.

\(~\text{typical dev./spread}\) (based on pop/dist).

Last time, we found the \(\mu_x (= E[x])\) of a bunch of famous distributions, and there is more in hw.
Now, let's find the \( \sigma_x^2 = V[x] \) of our special dist:

\[
\text{Binomial}(n, \pi): \quad V[x] = \sigma_x^2 = \sum_x (x - \mu_x)^2 \cdot p(x) = \cdots = n \pi (1-\pi).
\]

\[ \text{line} \quad \text{slope} = n \pi (1-\pi) \]

\[ \frac{1}{4n} \]

Parabola

Interpretation: \( \sigma_x \) is the (expected) typical deviation in \( x \).

So, \( \sigma_x \sim \sqrt{n} \)

The maximum \( \sigma_x \) is \( \frac{1}{2} \sqrt{n} \)

\[ \therefore \text{If you are tossing } n \text{ coins, the typical } X \text{ (# of heads)} \text{ is about } n \pi \text{ (i.e. } \mu_x \text{), and the typical dev. in } x \text{ is at most } \frac{1}{2} \sqrt{n}. \]

\[ \text{Poisson}(\lambda): \quad \sigma_x^2 = V[x] = \sum (x - 2)^2 \cdot p(x) = \cdots = 2 \]

\[ \text{Recall } E[x] = 2 < \quad \text{Same} \quad \]

\[ E[x] = \lambda \]

\[ \text{Normal}(\mu, \sigma): \quad \sigma_x^2 = V[x] = \int (x-\mu)^2 \cdot f(x) \, dx = \cdots = \sigma^2 \]

Which is why the poisson, \( \sigma^2 \) is called variance.

**Summary**

<table>
<thead>
<tr>
<th>( \mu_x )</th>
<th>( \sigma_x^2 )</th>
<th>( \mu )</th>
<th>( \sigma^2 )</th>
<th>Unit ((a, b))</th>
<th>( \frac{1}{\lambda} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n \pi )</td>
<td>( n \pi (1-\pi) )</td>
<td>( \lambda )</td>
<td>( \sigma^2 )</td>
<td>( (b-a)^2/12 )</td>
<td>( \frac{1}{\lambda^2} )</td>
</tr>
</tbody>
</table>
Jargon: histograms vs. distributions/pop.

Sample mean $\bar{x}$ vs. dist. mean $E[x] = \mu_x$

$\bar{x}$ variance $S^2$ vs. $\sigma^2$

Std. dev. $S$ vs. $\sigma_x$

One says that $\bar{x}$ is a point estimate of $\mu_x$, etc.

Called a "Statistic"  Called a "Parameter"}

Let's not forget areas, specifically $e.g.$ $1, 2, 1.96, \ldots$ (later)

The area within some std. dev. of the mean.

E.g. for $\mu(\mu, \sigma)$, 68% of the area is within 1 std. dev. of mean.

But now we can say things like this for any distr. e.g. Poisson, $

$\begin{array}{c}
\text{Computing areas like this will (eventually) allow us to provide some measure of confidence as to what $\mu_x$ is, based on observed data.}
\end{array}$

Recall,

For hist: area = prog. of times $x$ is observed to be within

For distr: area = prog. of times $x$ is expected to be within

[because we don't know the pop./dist! But if we assume the dist.]

That describes the pop., then we would expect ...$

Q1: For Exp(1), the area within 1 std. dev. of mean is equal to

a) 0  b) area between 0 and 1

c) area between 0 and $2/\lambda$  d) None of the above

$\mu_x - \sigma_x = \frac{1}{\lambda} - \frac{1}{\lambda} = 0$, $\mu_x + \sigma_x = 2/\lambda$
The business of estimating pop. params from sample stats refers to any distr. E.g., one says that $\bar{x}$ and $s$ provide point estimates of $\mu$ and $\sigma$ of of the normal dist. If the data come from a normal dist, to begin with.

Q: But, how do we know if our data come from a Normal dist?

Easier Q: How do we know if our data come from std. Normal?

A: Compare sample quantiles of data with distr. (or theoretical) quantiles.

Example: (Very crude!) Here is (sorted) data:

-1, 1, 2, 3, 4, 4.5, 5, 5.5, 6, 6.5, 8, 9

0.1 quantile: 0.1 quantile --- 0.5 quantile --- 0.9 quantile 1.0 quantile

→ I.e. The 0.1 sample quantile is 1, etc.

→ Theoretical quantiles: The 0.1 quantile of the std. normal, i.e.

$-\infty$ -1.285 $\cdots$ 0 $\cdots$ 1.285 $\infty$

$99.7\%$ plot: sample quantiles, replace with some large quantile, e.g., $-12.85, 1.285, 0.5, (0.5), (1.285, 8)$, replace with some small quantile, e.g., $-12.85, 1$. Theoretical quantiles.
If the histogram is consistent with a std. normal, then the quantiles/percentiles of data should be equal/compatible to those of the distr. Then the qq plot should be a straight diagonal line (intercept = 0, slope = 1).

If the data are not from std. normal, but from \( \mathcal{N}(\mu, \sigma) \), the only thing that changes is that the slope becomes \( \sigma \), and the intercept becomes \( \mu \). Not too obvious, but pfs in book.

In R: `qqnorm(x)` where \( x \) is the vector of data.

**Warning:**

On Fri I did not spend as much time on qq-plots as I would have liked. For that reason, I will keep talking about qq-plots on Mon. too. But qq-plots will be covered on the test on Fri.
From the histogram, it's hard to tell if the data come from a normal distribution, especially because histograms depend on bin size.

The plot looks linear, mostly! So, data are consistent with a normal. In fact, it looks like two different normals (bimodal) with different μ's, same σ (slope).
This exercise will help to get a better sense of what $\mu_x$ measures, geometrically.

Consider $f(x) = \begin{cases} 
1 + x & -1 < x \leq 0 \\
1 - x & 0 < x < 1 \\
0 & \text{else}
\end{cases}$

a) Plot (graph) $f(x)$ vs. $x$.

b) Confirm that $f(x)$ is a density function.

c) Compute the mean $\mu_x$.

d) Compute the variance $\sigma_x^2$.

Don't forget to check the solutions when it's posted to see an interpretation of $\sigma_x^2$.

For the uniform distr. (see 1.19) between $a, b$, show that

The expected value is $\frac{1}{2} (a + b)$, and the variance is $\frac{1}{12} (b - a)^2$.

Include these in your answers to hw-let9-2 and hw-let9-3.

For the exponential distr. with parameter $\lambda$, find $\mu_x$ and $\sigma_x^2$.

Hints: \[ \int_0^\infty ye^{-y} \, dy = 1 \quad \int_0^\infty (y-1)^2 e^{-y} \, dy = 1 \]

Find the area within one standard deviation of the mean (ie. $\mu_x \pm \sigma_x$) for

a) binomial $(n=20, \theta = \frac{1}{4})$

b) poisson $(\lambda = 5)$

c) Normal $(\mu = 5, \sigma = 1)$
In Example 1.23 (in text and in Lect), we found that on the average, out of 100 computers, 0.5 computers are defective.

a) What is the typical deviation we expect to see from this number (still out of 100)?

b) Suppose we do not know that the proportion of defective computers is 0.005. Then out of 100 computers, what is the maximum value we expect to see for typical deviation?