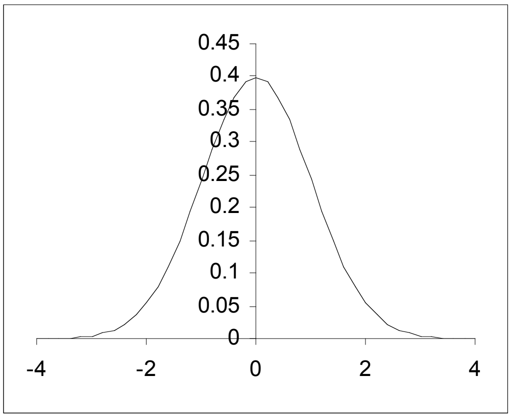
Lecture 5 ctd.
The Normal Distribution

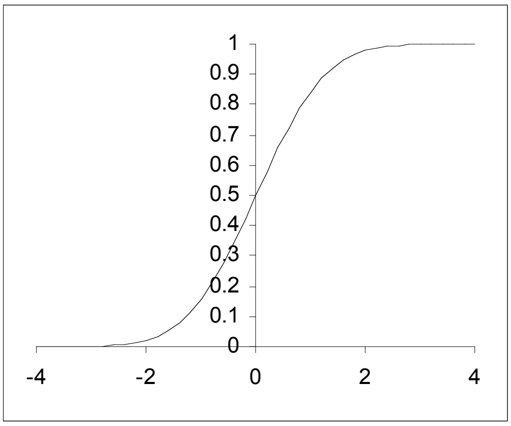
Also called the Gaussian distribution. This is probably the most fundamental model in statistics.

The model has two parameters: μ which is the mean and σ^2 which is the variance.

But we only need to describe the pdf for the standard version for which $\mu = 0$ and $\sigma^2 = 1$.

$$f_Z(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad F_X(z) = \Phi(z) = \int_{x=-\infty}^{x=z} \phi(x)dx$$





$NORMDIST(z, \mu, \sigma, 0)$ $\qquad$ $NORMDIST(z, \mu, \sigma, 1)$

If $\mu = 0$ and $\sigma = 1$ then cdf = $NORMSDIST(z)$

The Normal distribution plays an important role because it can be mathematically shown that sums (averages) of any random variables are approximately Normally distributed.
The result is known as the Central Limit Theorem.

It is the distribution of error – random variation about some value.
Errors are random and usually a sum of smaller errors.

Each VOL measurement is aiming at the same value – possibly the true value but not necessarily so.

This value is μ the mean parameter.

Because no two experiments are exactly identical (we would get the same value otherwise) we get slightly different errors of measurement. This leads to the variation. The extent of the variation is described by σ^2 the variance parameter.

Measurement Y = mean value + random variation
Leads to Normal distribution (at least approx).

The variation does not have to be error as such:

Consider the height of a male aged 19.

H = population mean + variation

The variation arises because of genetic differences, nutrition and possibly other things.

Again the height of males is well modelled by a Normal distribution.

The height of a person is not so well modelled because gender has an important impact on height. Height of females will also be Normal but with a different mean.

General Normals

A linear transformation of a Standard Normal is a general Normal.

$$Y = \mu + \sigma Z$$

$E(Y) = \mu$ and $V(Y) = \sigma^2$.

More usefully we transform the general Normal to a standard by:

$$Z = \frac{Y - \mu}{\sigma}$$

Calculating Normal Probabilities (for exams).

Standard Normal Probabilities ($\Phi(z)$) are tabulated in standard tables.

Thus: $P(Z \le a) = \Phi(a)$

$$P(a < Z \le b) = \Phi(b) - \Phi(a)$$

$$P(Z > a) = 1 - \Phi(a)$$

For a general Normal we transform to a standard Normal

$$Y \sim N(\mu, \sigma^2)$$

$$P(Y \le a) = P\left(\frac{Y-\mu}{\sigma} \le \frac{a-\mu}{\sigma}\right)$$

$$= P\left(Z \le \frac{a-\mu}{\sigma}\right) = \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$Y \sim Normal(10,25)$$

$$P(6 < Y \le 9) = P\left(\frac{6-10}{5} < Z \le \frac{9-10}{5}\right) = P(-0.8 < Z \le -0.2)$$

$$= \Phi(-0.2) - \Phi(-0.8) =$$

$$P(Y > 16) = P(Z > \frac{16-10}{5}) = P(Z > 1.2) = 1 - \Phi(1.2) =$$

Basic features:

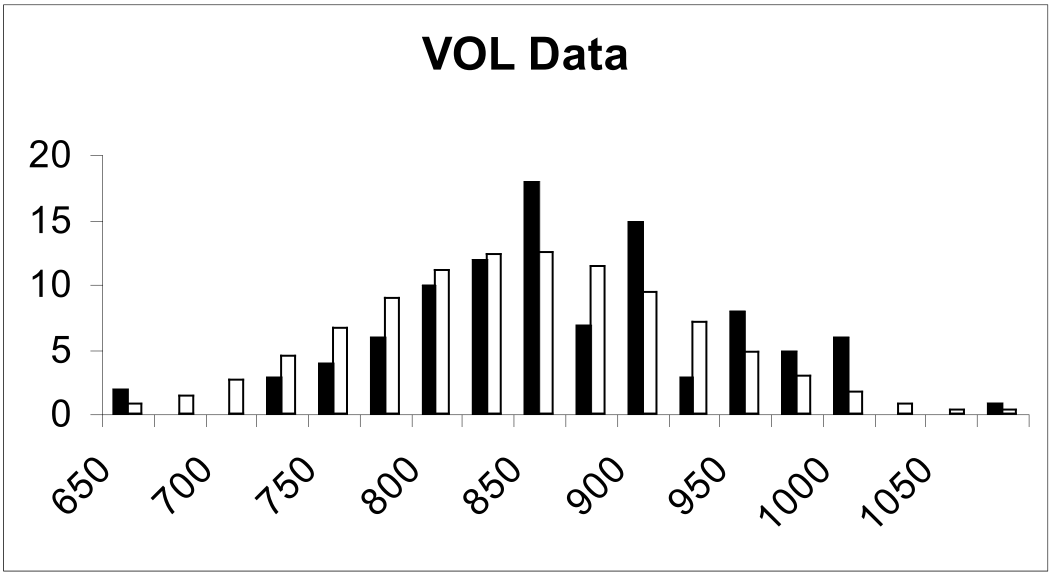
$$P(-1.96 < Z \le 1.96) = 0.95$$

1. 95% of the observations are within 2 (1.96) standard deviations of the mean.

2. 99.7% of the observations (practically all) are within 3 standard deviations of the mean.

Suppose we just take one observation and note that it is more than 2 standard deviations from the mean – we shall regard that as unusual.
If it is more than 3 standard deviations away this is extremely unusual.

VOL data



The solid black is the actual data frequency in the bin (650 – 699) (700-749) etc.
The white bars are the predicted frequencies for a

$$Normal(852.4,\ 79.01^2)$$

The Normal is probably a reasonable model for this data (later we shall see how to test this formally).
However the current estimates give a "true" value of 734.5 – the observed mean is a long way away from that, so there is a suspicion that their measurements were biased – their errors do not have 0 mean. (again a formal test of this later).

Standard Normal CDF										
x	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
-2.9	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001
-2.8	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
-2.7	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
-2.6	0.005	0.005	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
-2.5	0.006	0.006	0.006	0.006	0.006	0.005	0.005	0.005	0.005	0.005
-2.4	0.008	0.008	0.008	0.008	0.007	0.007	0.007	0.007	0.007	0.006
-2.3	0.011	0.010	0.010	0.010	0.010	0.009	0.009	0.009	0.009	0.008
-2.2	0.014	0.014	0.013	0.013	0.013	0.012	0.012	0.012	0.011	0.011
-2.1	0.018	0.017	0.017	0.017	0.016	0.016	0.015	0.015	0.015	0.014
-2	0.023	0.022	0.022	0.021	0.021	0.020	0.020	0.019	0.019	0.018
-1.9	0.029	0.028	0.027	0.027	0.026	0.026	0.025	0.024	0.024	0.023
-1.8	0.036	0.035	0.034	0.034	0.033	0.032	0.031	0.031	0.030	0.029
-1.7	0.045	0.044	0.043	0.042	0.041	0.040	0.039	0.038	0.038	0.037
-1.6	0.055	0.054	0.053	0.052	0.051	0.049	0.048	0.047	0.046	0.046
-1.5	0.067	0.066	0.064	0.063	0.062	0.061	0.059	0.058	0.057	0.056
-1.4	0.081	0.079	0.078	0.076	0.075	0.074	0.072	0.071	0.069	0.068
-1.3	0.097	0.095	0.093	0.092	0.090	0.089	0.087	0.085	0.084	0.082
-1.2	0.115	0.113	0.111	0.109	0.107	0.106	0.104	0.102	0.100	0.099
-1.1	0.136	0.133	0.131	0.129	0.127	0.125	0.123	0.121	0.119	0.117
-1	0.159	0.156	0.154	0.152	0.149	0.147	0.145	0.142	0.140	0.138
-0.9	0.184	0.181	0.179	0.176	0.174	0.171	0.169	0.166	0.164	0.161
-0.8	0.212	0.209	0.206	0.203	0.200	0.198	0.195	0.192	0.189	0.187
-0.7	0.242	0.239	0.236	0.233	0.230	0.227	0.224	0.221	0.218	0.215
-0.6	0.274	0.271	0.268	0.264	0.261	0.258	0.255	0.251	0.248	0.245
-0.5	0.309	0.305	0.302	0.298	0.295	0.291	0.288	0.284	0.281	0.278
-0.4	0.345	0.341	0.337	0.334	0.330	0.326	0.323	0.319	0.316	0.312
-0.3	0.382	0.378	0.374	0.371	0.367	0.363	0.359	0.356	0.352	0.348
-0.2	0.421	0.417	0.413	0.409	0.405	0.401	0.397	0.394	0.390	0.386
-0.1	0.460	0.456	0.452	0.448	0.444	0.440	0.436	0.433	0.429	0.425
0	0.500	0.496	0.492	0.488	0.484	0.480	0.476	0.472	0.468	0.464

x	0	1	2	3	4	5	6	7	8	9
0.0	0.500	0.504	0.508	0.512	0.516	0.520	0.524	0.528	0.532	0.536
0.1	0.540	0.544	0.548	0.552	0.556	0.560	0.564	0.567	0.571	0.575
0.2	0.579	0.583	0.587	0.591	0.595	0.599	0.603	0.606	0.610	0.614
0.3	0.618	0.622	0.626	0.629	0.633	0.637	0.641	0.644	0.648	0.652
0.4	0.655	0.659	0.663	0.666	0.670	0.674	0.677	0.681	0.684	0.688
0.5	0.691	0.695	0.698	0.702	0.705	0.709	0.712	0.716	0.719	0.722
0.6	0.726	0.729	0.732	0.736	0.739	0.742	0.745	0.749	0.752	0.755
0.7	0.758	0.761	0.764	0.767	0.770	0.773	0.776	0.779	0.782	0.785
0.8	0.788	0.791	0.794	0.797	0.800	0.802	0.805	0.808	0.811	0.813
0.9	0.816	0.819	0.821	0.824	0.826	0.829	0.831	0.834	0.836	0.839
1.0	0.841	0.844	0.846	0.848	0.851	0.853	0.855	0.858	0.860	0.862
1.1	0.864	0.867	0.869	0.871	0.873	0.875	0.877	0.879	0.881	0.883
1.2	0.885	0.887	0.889	0.891	0.893	0.894	0.896	0.898	0.900	0.901
1.3	0.903	0.905	0.907	0.908	0.910	0.911	0.913	0.915	0.916	0.918
1.4	0.919	0.921	0.922	0.924	0.925	0.926	0.928	0.929	0.931	0.932
1.5	0.933	0.934	0.936	0.937	0.938	0.939	0.941	0.942	0.943	0.944
1.6	0.945	0.946	0.947	0.948	0.949	0.951	0.952	0.953	0.954	0.954
1.7	0.955	0.956	0.957	0.958	0.959	0.960	0.961	0.962	0.962	0.963
1.8	0.964	0.965	0.966	0.966	0.967	0.968	0.969	0.969	0.970	0.971
1.9	0.971	0.972	0.973	0.973	0.974	0.974	0.975	0.976	0.976	0.977
2.0	0.977	0.978	0.978	0.979	0.979	0.980	0.980	0.981	0.981	0.982
2.1	0.982	0.983	0.983	0.983	0.984	0.984	0.985	0.985	0.985	0.986
2.2	0.986	0.986	0.987	0.987	0.987	0.988	0.988	0.988	0.989	0.989
2.3	0.989	0.990	0.990	0.990	0.990	0.991	0.991	0.991	0.991	0.992
2.4	0.992	0.992	0.992	0.992	0.993	0.993	0.993	0.993	0.993	0.994
2.5	0.994	0.994	0.994	0.994	0.994	0.995	0.995	0.995	0.995	0.995
2.6	0.995	0.995	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996
2.7	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997
2.8	0.997	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998
2.9	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.999	0.999	0.999