Fall 2003 EE290H Tentative Weekly Schedule

IC Yield & Performance

1. **Functional Yield of ICs and DFM.**
2. **Parametric Yield of ICs.**
3. **Yield Learning and Equipment Utilization.**

Process Modeling

4. **Statistical Estimation and Hypothesis Testing.** ←
5. **Analysis of Variance.**
6. **Two-level factorials and Fractional factorial Experiments.**

Process Control

7. **System Identification.**
8. **Parameter Estimation.**
9. **Statistical Process Control.** ***Distribution of projects. (week 9)***
10. **Run-to-run control.**
11. **Real-time control.** ***Quiz on Yield, Modeling and Control (week 12)***

Metrology

12. **Off-line metrology - CD-SEM, Ellipsometry, Scatterometry**
13. **In-situ metrology - temperature, reflectometry, spectroscopy**

Manufacturing Enterprise

14. **The Computer-Integrated Manufacturing Infrastructure**

15. ***Presentations of project results.***

A Brief Statistical Primer

Basic distributions. The central limit theorem. Sampling, estimation and hypothesis testing.

IC Production Requirements

- IC design must be modified for manufacturability.
- Incoming material and utilities must be qualified.
- Equipment must be qualified for volume production.
- Process must be debugged and transferred to high volume fab line.
- Yield detractors must be identified and eliminated.
- Efficient test procedures must be established.

Issues that SPC & DOE deal with

SPC helps remove the causes of problems and makes the process more stable over time.

DOE allows the exploration of the process settings and helps find the optimum operating conditions.

Quality is fitness for use.

There are two kinds of Quality:

a. Quality of design.

b. Quality of conformance.

DOE and SQC deal with quality of conformance.

SPC is one of the tools used to implement SQC.

Statistics is the art of making inferences about the whole by observing a sample.

The Evolution of Manufacturing Science

1Invention of machine tools. English system (1800).
mechanical - accuracy

2. Interchangeable components. American system (1850).
manufacturing - repeatability

3. Scientific management. Taylor system (1900).
industrial - reproducibility

4. Statistical Process Control (1930).
quality - stability

5. Information Processing and Numerical Control (1970).
system - adaptability

6. Intelligent Systems and CIM (1980).
knowledge – versatility

7. Physical and logical (“lights out”) Automation (2000).
integration – efficiency

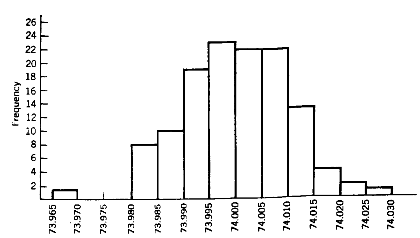
Outline

- Discrete Distributions
- Continuous Distributions (Normal)
- Central Limit Theorem

- Sampling
- Estimation
- Hypothesis Testing

Presenting and Summarizing Data

The histogram



The sample average

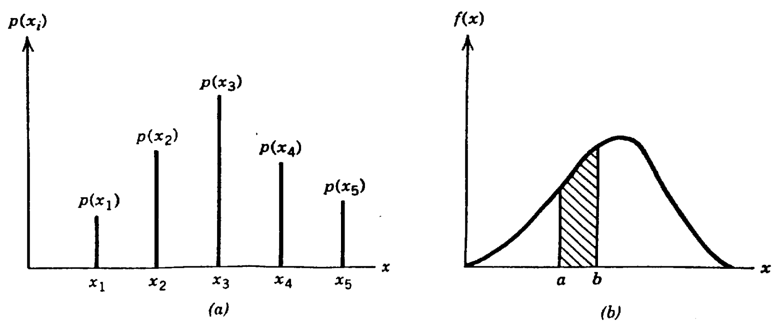
$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

The sample variance

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

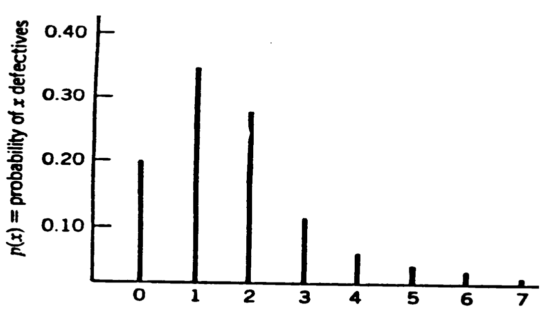
The Distribution of x

Discrete Distributions Continuous Distributions



Probability distributions. (*a*) Discrete case. (*b*) Continuous case.

A discrete distribution: the binomial



Binomial Distribution with p = 0.10 and n=15.

$$P\{D = x\} = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0,1,...,n$$

$$\text{mean} = np, \ \text{variance} = np(1-p)$$

$$\text{where} \binom{n}{x} \equiv \frac{n!}{x!(n-x)!}, \quad n! \equiv 1 \cdot 2 \cdot ... \cdot (n-1) \cdot n$$

$$\text{the sample fraction} \quad \hat{p} = \frac{D}{n} \qquad \text{mean} = p, \quad \text{variance} = \frac{p(1-p)}{n}$$

Example: Using the Binomial Distribution

The probability of a broken wafer in a loading operation is 0.1% After loading 100 wafers, what is the probability that:

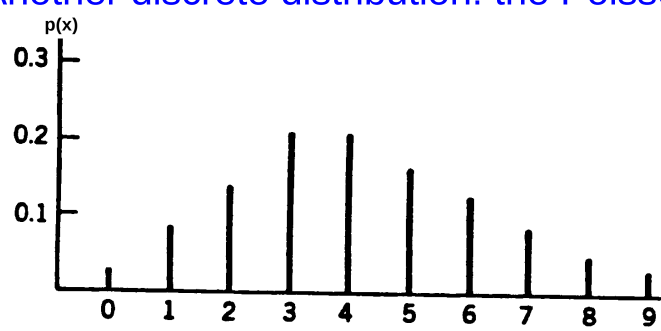
All wafers are OK?

1 wafer broke?

5 wafers broke?

More than 5 wafers broke?

Another discrete distribution: the Poisson



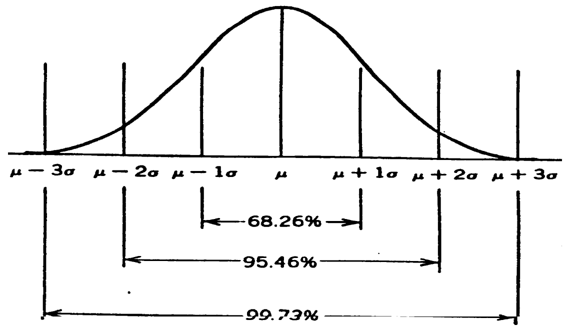
The Poisson Distribution with l = 4.

$$p(x) = \frac{e^{-\lambda} \lambda^x}{x!} \qquad x = 0, 1, .. \qquad \text{where } x! \equiv 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (x-1) \cdot x$$

mean: $\mu = \lambda$

variance: $\sigma^2 = \lambda$

A continuous distribution: the Normal



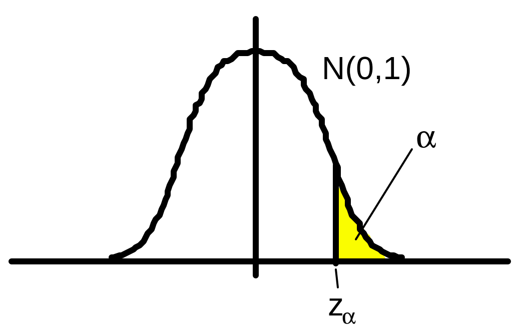
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty \qquad x \sim N(\mu, \sigma^2)$$

$$P\{x < a\} = F(a) = \int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$z \equiv \frac{x-\mu}{\sigma}, \quad P\{x < a\} = P\left\{z < \frac{a-\mu}{\sigma}\right\} \equiv \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Useful Definitions on the Normal Distribution

$z \sim N(0, 1)$ This is the Standard Normal Distribution



for any Normal distribution, z signifies the "number of sigmas away from the mean".
z_α is the number that "cuts" an upper tail with area α

Notice that $z_\alpha = -z_{1-\alpha}$

The Normal Distribution

THE
NORMAL
LAW OF ERROR
STANDS OUT IN THE
EXPERIENCE OF MANKIND
AS ONE OF THE BROADEST
GENERALIZATIONS OF NATURAL
PHILOSOPHY ♦ IT SERVES AS THE
GUIDING INSTRUMENT IN RESEARCHES
IN THE PHYSICAL AND SOCIAL SCIENCES AND
IN MEDICINE AGRICULTURE AND ENGINEERING ♦
IT IS AN INDISPENSABLE TOOL FOR THE ANALYSIS AND THE
INTERPRETATION OF THE BASIC DATA OBTAINED BY OBSERVATION AND EXPERIMENT

Example: Table Lookup for Normal Distribution

The wafer to wafer thickness of a poly layer is distributed normally around 500nm with a σ of 20nm:

$$Pth \sim N(500nm, 400nm^2)$$

What is the probability that a given wafer will have polysilicon thicker than 510nm?

... thinner than 480nm?

... between 490 and 515nm?

The Additivity of Variance

If $y = a_1x_1+a_2x_2+...+ a_nx_n$ then $\mu_y = a_1\mu_1+...+ a_n\mu_n$
and
$$\sigma_y^2 = a_1^2\sigma_1^2+...+a_n^2\sigma_n^2$$

This applies under the assumption that the parameters x are *independent*..

For example, consider the thickness variance of a layer defined by two consecutive growths:

$$\mu_t = \mu_{g1}+\mu_{g2}$$
$$\sigma_t^2 = \sigma_{g1}^2+ \sigma_{g2}^2$$

or by a growth step followed by an etch step:

$$\mu_t = \mu_g-\mu_e$$
$$\sigma_t^2 = \sigma_g^2+ \sigma_e^2$$

Example: How to combine consecutive steps

The thickness of a SiO_2 layer is distributed normally around 600nm with a σ of 20nm:

GOx~N (600nm, 400nm^2)

During a polysilicon removal step with limited selectivity, some of the oxide is removed. The removed oxide is:

ROx~N (50nm, 25nm^2)

What is the probability that the final oxide thickness is between 540 and 560nm?

The Correlation Coefficient

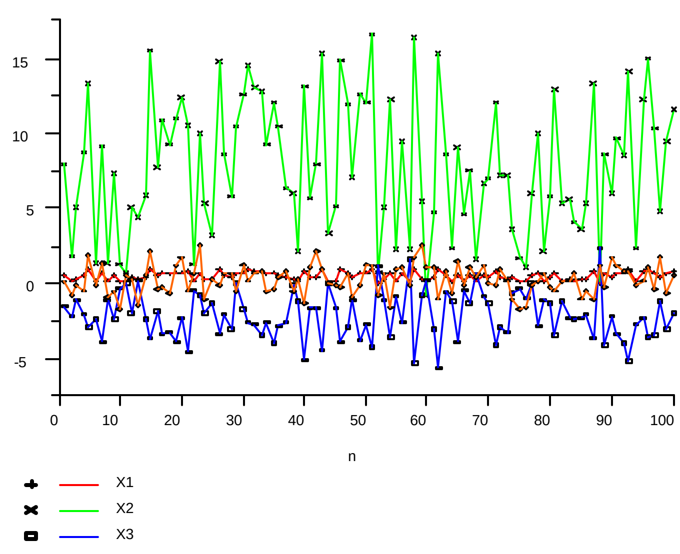
The correlation coefficient ρ, is a statistical moment that gives a measure of *linear dependence* between two random variables. It is *estimated* by:

$$r = \frac{s_{xy}}{s_x s_y}$$

where s_x and s_y are the square roots of the estimates of the variance of x and y, while s_{xy} is an estimate of the *covariance* of the two variables and is estimated by:

$$s_{xy} = \sum_{i=1}^{n} \frac{(x_i - \overline{x})(y_i - \overline{y})}{n-1}$$

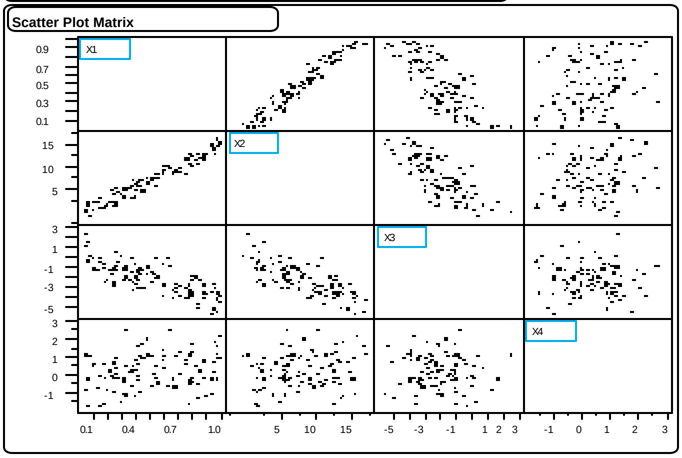
Correlation Example



Correlation Example

Correlations

Variable	X1	X2	X3	X4
X1	1.0000	0.9719	-0.7728	0.2089
X2	0.9719	1.0000	-0.7518	0.2061
X3	-0.7728	-0.7518	1.0000	-0.0753
X4	0.2089	0.2061	-0.0753	1.0000



Using the Correlation Coefficient

In fact, the only way to stop variances from accumulating is by making sure that the parameters have the appropriate *correlation* ρ.

$$\mu_z = a\mu_x + b\mu_y$$

$$\sigma_z^2 = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\rho\,\sigma_x\sigma_y$$

For example, two consecutive growths can be controlled if we set $\rho = -1$

$$\mu_t = \mu_{g1} + \mu_{g2}$$

$$\sigma_t^2 = \sigma_{g1}^2 + \sigma_{g2}^2 - 2\sigma_{g1}\sigma_{g2}$$

and a growth followed by an etch can be controlled by setting $\rho = 1$

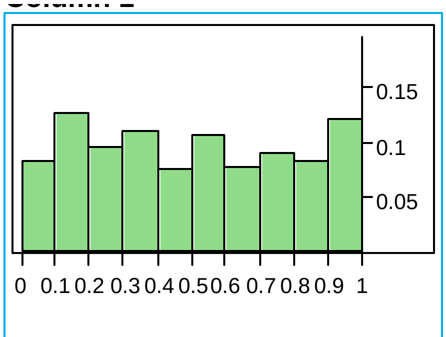
$$\mu_t = \mu_g - \mu_e$$

$$\sigma_t^2 = \sigma_g^2 + \sigma_e^2 - 2\sigma_g\sigma_e$$

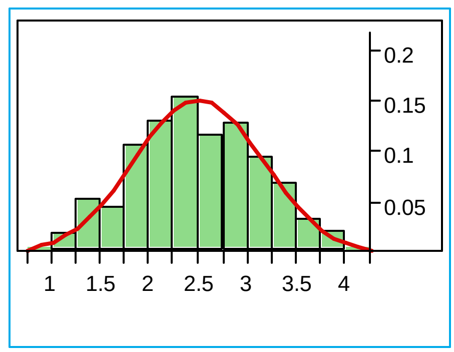
The Central Limit Theorem:

The sum of <u>independent</u> random variables <u>tends</u> to have a normal distribution.

Uniformly distributed number



Sum of 5 unif. distr. numbers:



Statistical Moments and their Estimators

The mean (μ), the sigma (σ), the correlation (ρ), etc. are symbolized with Greek characters and they refer to "true", but hidden values *which we cannot know exactly*.

These are the *moments* of a population.

The average, the standard deviation, the correlation coefficient, etc. are represented by Latin characters and they refer to values, *which we can calculate from the data.*

These values are used to estimate the true (but impossible to know exactly) moments. They are *estimators*.

Sampling and Estimation

Sampling: the act of making inferences about populations.

Random sampling: when each observation is identically and independently distributed.

Statistic: a function of sample data containing no unknowns. (e.g. average, median, standard deviation, etc.)

A statistic is a random variable. Its distribution is a sampling distribution.

EE290H F03 Spanos & Poolla

Example: Estimating the mean of a normal dist.

The thickness of a poly layer is distributed normally around 500nm with a σ of 10nm:

$$P_{th} \sim N(500\text{nm}, 100\text{nm}^2)$$

We randomly select 50 wafers, measure the poly thickness and calculate the average of the fifty readings:

$$\overline{P_{th}} = \frac{1}{50}\sum_{i=1}^{50} P_{th\,i}$$

What is the distribution of $\overline{P_{th}}$?

What is the probability that $\overline{P_{th}}$ will be between 495 and 505nm?

Example: The Mean of a Normal Dist. (cont)

Now that we understand how the average is distributed, lets use it to estimate the mean (assuming that we know that σ is 10nm).

If the 50 measurements yielded an average of 503nm, what can we say about the unknown mean?

What is the estimated value of the unknown mean?

What is the probability that the unknown mean is between 500 and 506nm?

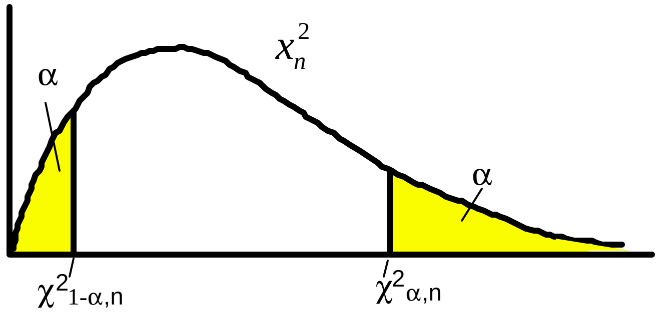
The chi-square distribution

The *chi-square (c^2)* distribution with n degrees of freedom is defined as:

$$\text{if } x_i \sim N(0,1) \text{ for } i = 1,2,...,n \text{ then:}$$
$$x_1^2 + x_2^2 + ... + x_n^2 \sim \chi^2_n$$

Usage: Defines the distribution of the standard deviation when the mean is known. Also to define the student-t distribution (more on this later).

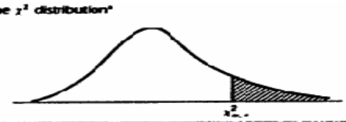
Useful Definitions around the χ^2 distribution



Statistical tables can give us the value of the random number that "cuts" an α upper tail of the distribution.

The χ^2 distribution Table

Percentage points of the χ^2 distribution*



ν	0.995	0.990	0.975	0.950	0.500	0.050	0.025	0.010	0.005
1	0.00+	0.00+	0.00+	0.00+	0.45	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	1.39	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	2.37	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	3.36	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	4.35	11.07	12.38	15.09	16.75
6	0.68	0.87	1.24	1.64	5.35	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	6.35	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	7.34	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	8.34	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	9.34	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	10.34	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	11.34	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	12.34	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	13.34	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	14.34	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	15.34	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	16.34	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	17.34	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	18.34	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	19.34	31.41	34.17	37.57	40.00
25	10.52	11.52	13.12	14.61	24.34	37.65	40.65	44.31	46.93
30	13.79	14.95	16.79	18.49	29.34	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	39.34	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	49.33	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	59.33	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	69.33	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	79.33	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	89.33	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	99.33	124.34	129.56	135.81	140.17

ν = degrees of freedom.

* Adapted with permission from *Biometrika Tables for Statisticians*, Vol. 1, 3rd ed., by E. S. Pearson and H. O. Hartley, Cambridge University Press, Cambridge, 1966.

The *student-t* distribution

The *student-t* distribution with k degrees of freedom is defined as:

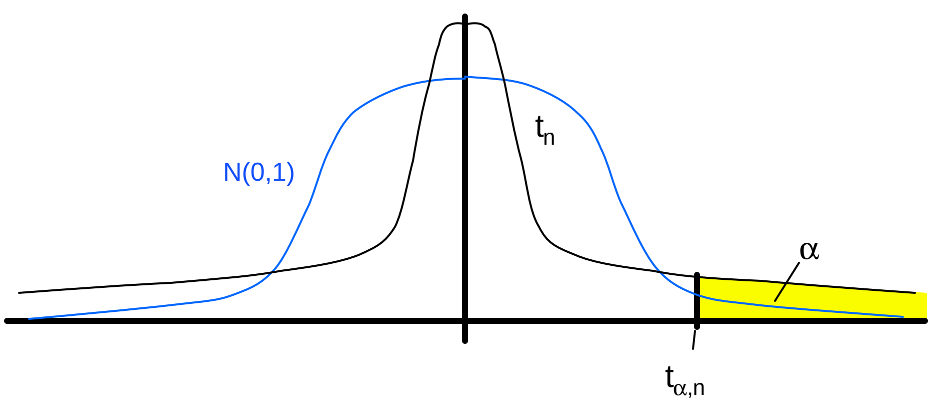
if $z \sim N(0,1)$ then:

$$\frac{z}{\sqrt{y/n}} \sim t_n$$

$$y \sim \chi^2_n$$

<u>Usage</u>: Find the distribution of the average when σ is unknown.

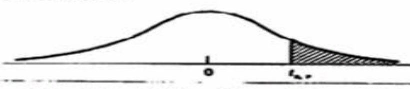
Useful Definitions about the *student-t*



Again, $t_{\alpha,n}$ signifies the value of the random number that cuts an a upper tail with area "α".
Notice that $t_{\alpha,n} = - t_{1-\alpha,n}$.
Also notice that $t_{\alpha,n}$ is a bit larger than z_α.

Student-t Distribution table

Percentage points of the t distribution*



ν \ α	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.265	0.727	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.019	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	0.254	0.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

ν = degrees of freedom.

* Adapted with permission from *Biometrika Tables for Statisticians*, Vol. 1, 3rd ed., by E.S. Pearson and H.O. Hartley, Cambridge University Press, Cambridge, 1966.

The F distribution

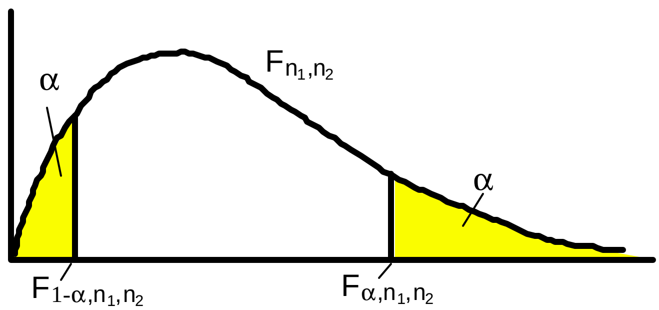
The F distribution with n_1, n_2 degrees of freedom is defined as:

$$\frac{y_1/n_1}{y_2/n_2} \sim F_{n_1,n_2}$$

$$y_1 \sim \chi^2_{n_1} \qquad y_2 \sim \chi^2_{n_2}$$

Usage: to compare the spread of two populations.

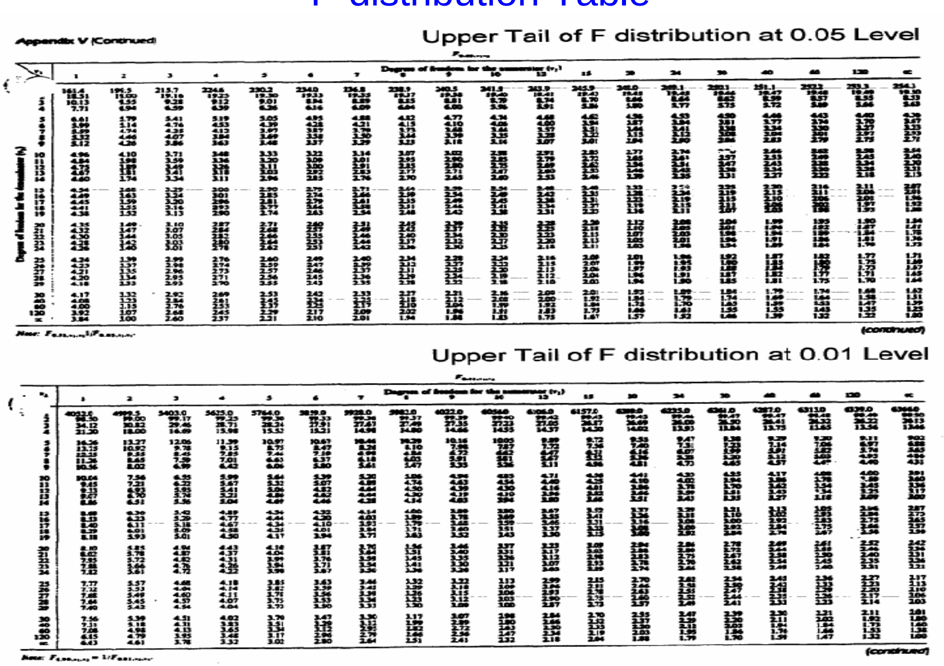
Useful Definitions around the F distribution



Statistical tables can give us the value of the random number that “cuts” an α upper tail of the F distribution.

It so happens that $F_{1-\alpha,n_1,n_2} = \frac{1}{F_{\alpha,n_1,n_2}}$

F-distribution Table



Sampling from a Normal Distribution

The statistics of the average:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \ \sigma \text{ is known}$$

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t_{n-1}, \ \sigma \text{ is } \textit{not} \text{ known}$$

The statistics of the standard deviation:

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}$$

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1} \qquad \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim F_{n_1-1, n_2-1}$$

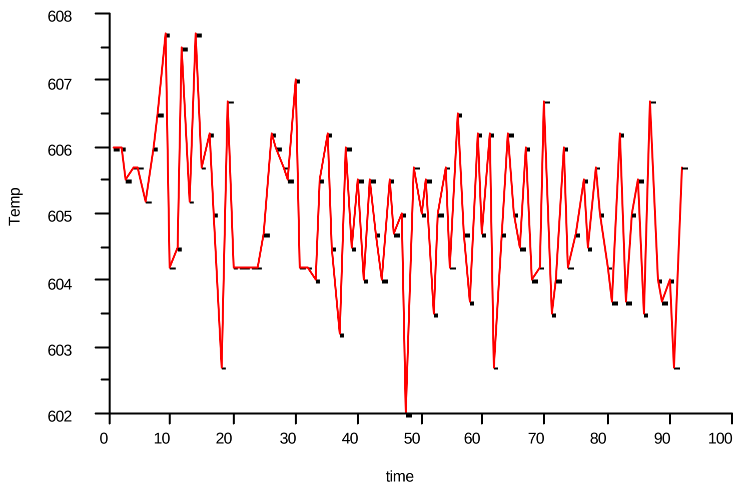
Estimation: The art of guessing at unknowns

Point estimation must be unbiased and must have minimum variance.

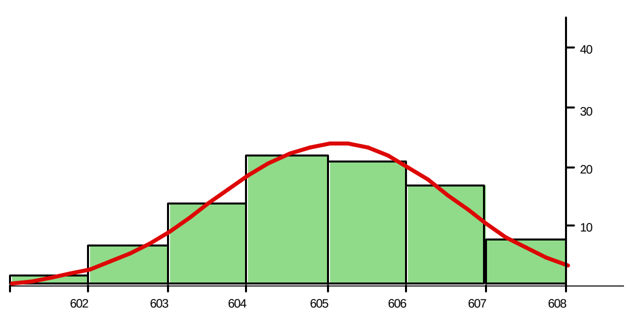
Interval estimation yields bounds that contain the actual value with a given certainty.

(For interval estimation we need the sampling distribution.)

Example - Analyze LPCVD Temp. Readings



Example: Analyzing Temperature Readings



Moments

Mean	605.0791
Std Dev	1.5058
Std Err Mean	0.1579
upper 95% Mean	605.3927
lower 95% Mean	604.7655
N	91.0000
Sum Wgts	91.0000

Example: The 95% confidence intervals on furnace temperature statistics.

Mean:

$$\bar{x} - t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}$$

or, at the 5% level, $\mu = 605.1 \pm 0.313$

Variance:

$$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}$$

or, at the 5% level, $1.31 < \sigma < 1.76$

Example: Use the tables to find the 1% confidence intervals for μ and σ, when:

- Average: 605.1μm
- Standard Deviation: 1.51μm
- N=91
- (The distribution is assumed to be normal.)

Hypothesis Testing

Hypothesis testing is a selection between:

Ho (null hypothesis)

and

H1 (alternative hypothesis).

α = P{type I error} = P{reject Ho / Ho is true}

β = P{type II error} = P{accept Ho / Ho is false}

Power = 1 - β = P{reject Ho / Ho is false}

<u>Example</u>:

Ho : the means of two populations are equal.

H1 : the means of two populations are different.

Comparing Population Mean to a Constant

(to simplify this, we assume that the variances are known)

Since the arithmetic average is distributed as:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

given a constant μ_o , we have:

$$\bar{x} - \mu_o \sim N\left(\mu - \mu_o, \frac{\sigma^2}{n}\right)$$

and under the H_o that $\mu = \mu_o$, we have:

$$\bar{x} - \mu_o \sim N\left(0, \frac{\sigma^2}{n}\right)$$

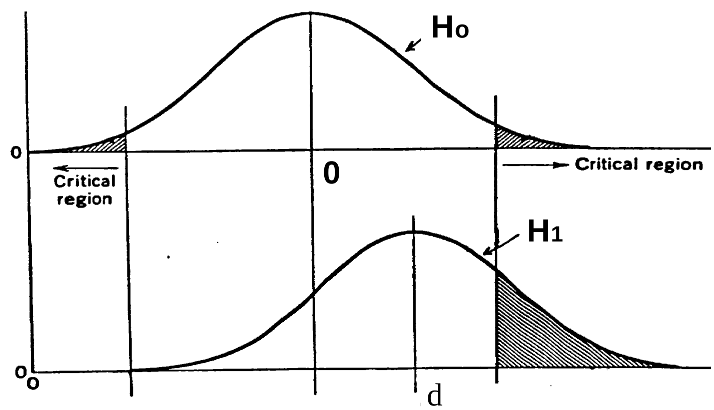
Under this assumption, we can find the probability that $\bar{x} - \mu_o$ will differ from zero by a certain amount...

Comparing Mean to a Constant (cont)

If the observed value of $\bar{x} - \mu_o$ is so far from zero that its probability under H_0 would be small, the H_0 should be rejected.

H_o is rejected at a *level of significance* "α" when $\bar{x} - \mu_o$ falls in a predesignated area of the H_0 distribution that is selected to correspond to a small probability "α".

Graphically:



Example:
Is the mean of the temperature 605 degrees?

Data File: tyl16stats L1

Single Sample...

Variable:	center	Population
Mean:	6.0510e+2	6.0500e+2
Std. Deviation:	1.5108e+0	
Observations:	92	

t-statistic:	6.3486e-1	Hypothesis:
Degrees of Freedom:	91	Ho: µ1 = µ2
Significance:	5.27e-1	Ha: µ1 ° µ2

Use the student-t tables to test the hypothesis:

H_0: The mean temperature is 605 degrees.

H_1: The mean temperature is *not* 605 degrees.

Step 1: Collect your data.

Step 2: Assume that H_0 is true.

Step 3: Calculate the appropriate test statistic.

Step 4: Place the statistic against its assumed distribution and see how well it fits.

Step 5: Accept or reject H_0 based on Step 4.

More Tests

Simple tests can also be devised to compare the mean value to a standard when the sigma is unknown.

Simple tests are also in use to compare two (or more) population means with sigmas known, unknown, identical or not.

A test based on the F distribution is available to compare sigmas.

Tests are also available to compare correlation coefficients, etc.

These tests are implicitly or explicitly used in experimental design when we compare the results of various treatments.

More Tests to Compare Populations

Compare means if the two sigmas are the same...

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$$

$$\frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

Compare means if the two sigmas are different...

$$n = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1+1}+\frac{(s_2^2/n_2)^2}{n_2+1}}$$

$$\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \sim t_n$$

Summary

Distributions are used to describe properties of populations.

Moments are used to describe aspects of distributions.

Estimators are used to approximate the values of moments.

Moments can be estimated as numbers and/or as intervals.

Populations can be compared by building and testing various Hypotheses about their moments.

What does all this have to do with the manufacturing of ICs?