4. Variance analysis & regression analysis

- Different distributions of random variables
- Confidence intervals
- Hypothesis tests (statistical tests)
- Variance analysis
- Regression analysis (least squares fitting)

Statiscal properties of a random variable ε

- Distribution function $F(x)$
  probability that $\varepsilon < x$
  $$P(\varepsilon \leq x) = F(x) = \int_{-\infty}^{x} f(\varepsilon) d\varepsilon \qquad (-\infty < x < +\infty)$$
- Density function $f(x)$
  $$\frac{\partial F(x)}{\partial x} = f(x)$$
- Expectation (*average*)
  $$E(\varepsilon) = \int_{-\infty}^{+\infty} \varepsilon \, f(\varepsilon) d\varepsilon$$
- Variance
  $$var(\varepsilon) = \sigma^2 = E\left\{[\varepsilon - E(\varepsilon)]^2\right\} = \int_{-\infty}^{+\infty} [\varepsilon - E(\varepsilon)]^2 \cdot f(\varepsilon) \cdot d\varepsilon$$

Several types of distributions

- Binomial distribution
- Uniform distribution
- Normal distribution.
  Standard normal distribution
- χ^2-distribution
- t-distribution
- F-distribution

Binomial distribution

- A random experiment with only two outcomes
- Probabilities: p, q, $p + q = 1$
- Density (frequency) function: Probability that experiment is repeated n times and outcome with probability p occurs k times

$$f(k) = \binom{n}{k} p^k q^{n-k} = \frac{n!}{(n-k)!\,k!} p^k q^{n-k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k(k-1)(k-2)\cdots 3\cdot 2\cdot 1} p^k q^{n-k} \quad (k = 0, 1, 2, \cdots, n)$$

- Distribution function: $F(x) = \sum_{k=0}^{x} f(k) = \sum_{k=0}^{x} \binom{n}{k} p^k q^{n-k}$
- Expectation and variance:

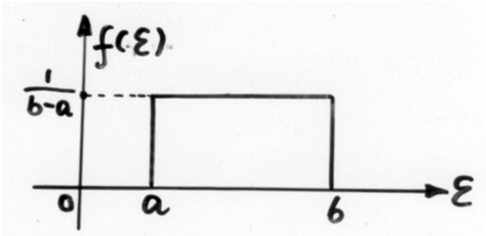
$$E(x) = \sum_{k=0}^{x} k f(k) = np\ , \quad Var(x) = \sum_{k=0}^{x} \left\{ (k-np)^2 f(k) \right\} = npq$$

Example
the sign of errors:
p = q = ½

Uniform distribution

- Constant density function inside interval

$$f(\varepsilon) = \begin{cases} \frac{1}{b-a} & \text{for } a \le \varepsilon \le b \\ 0 & \text{otherwise} \end{cases}$$



- Distribution function:

$$F(x) = \begin{cases} 0 & \text{for } x < a \\ \frac{x-a}{b-a} & \text{for } a \le x \le b \\ 1 & \text{for } x > b \end{cases}$$

Uniform distribution

- Expectation, Variance:

$$E(\varepsilon) = \int_a^b \varepsilon \frac{1}{b-a} d\varepsilon = \frac{a+b}{2}$$

$$Var(\varepsilon) = E\left\{[\varepsilon - E(\varepsilon)]^2\right\} = \frac{(b-a)^2}{12}$$

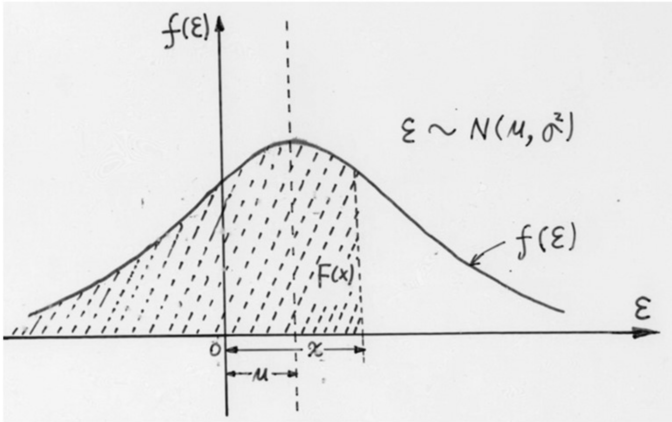
Example

Rounding errors: a= -½, b= +½

Expectation and variance: $E(\varepsilon) = 0, \quad Var(\varepsilon) = \frac{1}{12}$

Density function of a normal distribution

$$f(\varepsilon) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(\varepsilon-\mu)^2\right\} \qquad (-\infty < \varepsilon < +\infty)$$



$$F(x) = \int_{-\infty}^{x} f(\varepsilon)d\varepsilon = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{1}{2\sigma^2}(\varepsilon-\mu)^2\right\} \cdot d\varepsilon \quad (-\infty < x < +\infty)$$

Normal distribution $\varepsilon \sim N(\mu, \sigma^2)$

- Expectation:

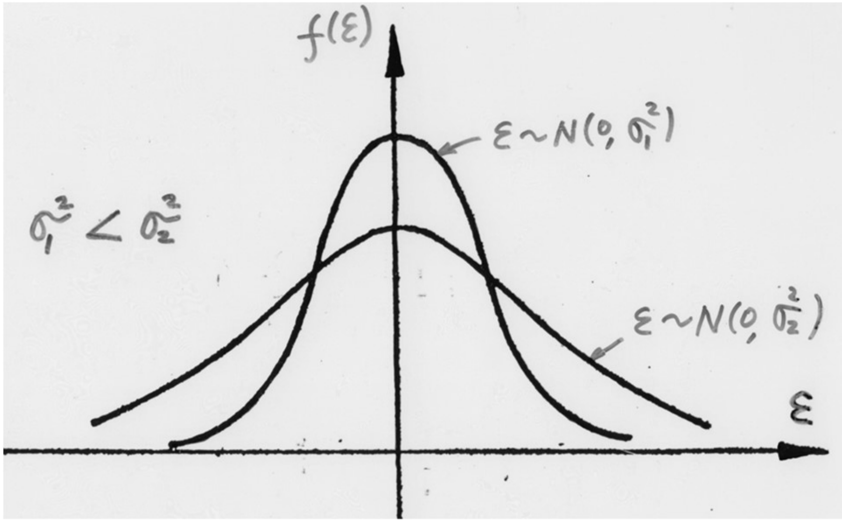
$$E(\varepsilon) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varepsilon \cdot \exp\left\{-\frac{1}{2\sigma^2}(\varepsilon-\mu)^2\right\} \cdot d\varepsilon = \mu$$

- Variance:

$$Var(\varepsilon) = E\left\{[\varepsilon - E(\varepsilon)]^2\right\} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} [\varepsilon - E(\varepsilon)]^2 \cdot \exp\left\{-\frac{1}{2\sigma^2}(\varepsilon-\mu)^2\right\} \cdot d\varepsilon = \sigma^2$$

$\longrightarrow$
- Normal distribution is defined by (μ, σ^2)
- Most errors follow normal distributions
- *Central limiting theorem: Sum of n independent random variables of equal expectations/variances converges to a normal distribution*

Two normal distributions of zero means



Standard normal distribution $N(0,1)$

$$\varepsilon \sim N(\mu, \sigma^2) \qquad \tau = \frac{\varepsilon - \mu}{\sigma} \quad \longrightarrow \quad E(\tau) = 0\,, \quad Var(\tau) = 1$$

$\longrightarrow$ τ is said to have *standard normal distribution*

$$\tau \sim N(0,1)$$

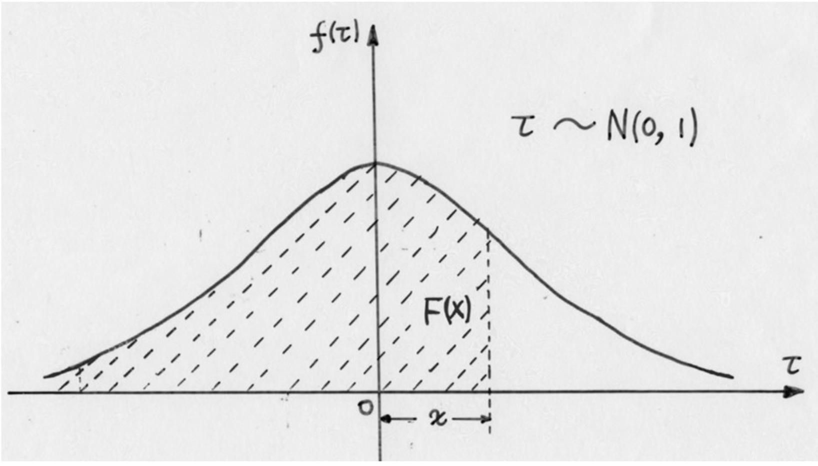
- Density function:

$$f(\tau) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\tau^2\right\} \qquad (-\infty < \tau < +\infty)$$

- Distribution function:

$$F(x) = \int_{-\infty}^{x} f(\tau)d\tau = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{1}{2}\tau^2\right\} \cdot d\tau \quad (-\infty < x < +\infty)$$

Standard normal distribution table



$x =$	0	1	2	3
$F(x) =$	0.5	0.8413	0.97725	0.99865

Standard normal distribution table

$$F(x) = P(\varepsilon \leqslant x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{\varepsilon^2}{2}} \cdot d\varepsilon$$

x	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9648	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Normal distribution of multiple variables

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \varepsilon_n \end{bmatrix}$$

$$E(\varepsilon) = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \cdots \\ \mu_n \end{bmatrix} = \underset{n\cdot 1}{\mu}, \quad \underset{n\cdot n}{C} = E\left\{[\varepsilon - E(\varepsilon)]\,[\varepsilon - E(\varepsilon)]^\top\right\} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix}$$

$$f(\varepsilon) = \frac{1}{(2\pi)^{\frac{n}{2}}\,|C|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\,(\varepsilon - \mu)^\top C^{-1}\,(\varepsilon - \mu)\right\}$$

χ^2-distribution

$$\chi^2 = \varepsilon_1^2 + \varepsilon_2^2 + \cdots + \varepsilon_n^2 \qquad \varepsilon_i \sim N(0,1)$$

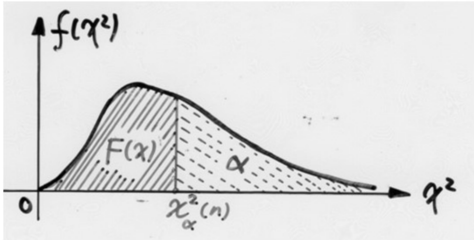
n is called the degree of freedom of χ^2

$$f(\chi^2) = \frac{1}{\frac{n}{2}\Gamma(\frac{n}{2})}\ \left(\chi^2\right)^{\frac{n}{2}-1} \exp\left\{-\frac{1}{2}\chi^2\right\}, \quad (\chi^2 > 0)$$

$$\Gamma(n) = \int_0^{+\infty} x^{n-1} e^{-x} dx \quad (n > 0)$$

$$F(x) = \int_{-\infty}^{x} f(x) dx$$

$$E\left(\chi^2\right) = n\,, \quad Var\left(\chi^2\right) = E\left\{\left[\chi^2 - E(\chi^2)\right]^2\right\} = 2n$$



n	α=0.995	0.99	0.975	0.95	0.90	0.75
1	—	—	0.001	0.004	0.016	0.102
2	0.010	0.020	0.051	0.103	0.211	0.575
3	0.072	0.115	0.216	0.352	0.584	1.213
4	0.207	0.297	0.484	0.711	1.064	1.923
5	0.412	0.554	0.831	1.145	1.610	2.675
6	0.676	0.872	1.237	1.635	2.204	3.455
7	0.989	1.239	1.690	2.167	2.833	4.255
8	1.344	1.646	2.180	2.733	3.490	5.071
9	1.735	2.088	2.700	3.325	4.168	5.899
10	2.156	2.558	3.247	3.940	4.865	6.737
11	2.603	3.053	3.816	4.575	5.578	7.584
12	3.074	3.571	4.404	5.226	6.304	8.438
13	3.565	4.107	5.009	5.892	7.042	9.299
14	4.075	4.660	5.629	6.571	7.790	10.165
15	4.601	5.229	6.262	7.261	8.547	11.037
16	5.142	5.812	6.908	7.962	9.312	11.912
17	5.697	6.408	7.564	8.672	10.085	12.792
18	6.265	7.015	8.231	9.390	10.865	13.675
19	6.844	7.633	8.907	10.117	11.651	14.562
20	7.434	8.260	9.591	10.851	12.443	15.452
21	8.034	8.897	10.283	11.591	13.240	16.344
22	8.643	9.542	10.982	12.338	14.042	17.240
23	9.260	10.196	11.689	13.091	14.848	18.137
24	9.886	10.856	12.401	13.848	15.659	19.037
25	10.520	11.524	13.120	14.611	16.473	19.939
26	11.160	12.198	13.844	15.379	17.292	20.843
27	11.808	12.879	14.573	16.151	18.114	21.749
28	12.461	13.565	15.308	16.928	18.939	22.657
29	13.121	14.257	16.047	17.708	19.768	23.567

Note the risk level α and the correpsonding critical value

t-distribution, *t(n)*

If $X \sim N(0,1)$ and $Y \sim \chi^2(n)$ are independent of each other, then

$$t = \frac{X}{\sqrt{Y/n}}$$

follows the t-distribution with degree of freedom n

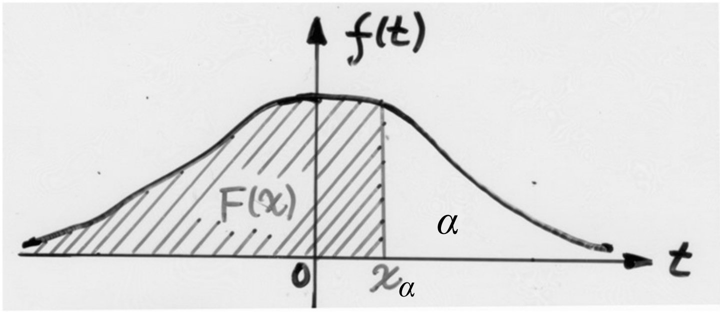
$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \cdot \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \qquad (-\infty < t < +\infty)$$

$$E(t) = 0 \;\; (n > 1), \qquad Var(t) = \frac{n}{n-2} \;\; (n > 2)$$

t(n) approaches $N(0,1)$ when n approaches infinity

Curve of the t-distribution

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \cdot \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \qquad (-\infty < t < +\infty)$$



The t-distribution table

n	α = 0.995	0.99	0.975	0.95	0.90	0.75
1	—	—	0.001	0.004	0.016	0.102
2	0.010	0.020	0.051	0.103	0.211	0.575
3	0.072	0.115	0.216	0.352	0.584	1.213
4	0.207	0.297	0.484	0.711	1.064	1.923
5	0.412	0.554	0.831	1.145	1.610	2.675
6	0.676	0.872	1.237	1.635	2.204	3.455
7	0.989	1.239	1.690	2.167	2.833	4.255
8	1.344	1.646	2.180	2.733	3.490	5.071
9	1.735	2.088	2.700	3.325	4.168	5.899
10	2.156	2.558	3.247	3.940	4.865	6.737
11	2.603	3.053	3.816	4.575	5.578	7.584
12	3.074	3.571	4.404	5.226	6.304	8.438
13	3.565	4.107	5.009	5.892	7.042	9.299
14	4.075	4.660	5.629	6.571	7.790	10.165
15	4.601	5.229	6.262	7.261	8.547	11.037
16	5.142	5.812	6.908	7.962	9.312	11.912
17	5.697	6.408	7.564	8.672	10.085	12.792
18	6.265	7.015	8.231	9.390	10.865	13.675
19	6.844	7.633	8.907	10.117	11.651	14.562
20	7.434	8.260	9.591	10.851	12.443	15.452
21	8.034	8.897	10.283	11.591	13.240	16.344
22	8.643	9.542	10.982	12.338	14.042	17.240
23	9.260	10.196	11.689	13.091	14.848	18.137
24	9.886	10.856	12.401	13.848	15.659	19.037
25	10.520	11.524	13.120	14.611	16.473	19.939
26	11.160	12.198	13.844	15.379	17.292	20.843
27	11.808	12.879	14.573	16.151	18.114	21.749
28	12.461	13.565	15.308	16.928	18.939	22.657
29	13.121	14.257	16.047	17.708	19.768	23.567
30	13.787	14.954	16.791	18.493	20.599	24.478
31	14.458	15.655	17.539	19.281	21.434	25.390
32	15.134	16.362	18.291	20.072	22.271	26.304
33	15.815	17.074	19.047	20.867	23.110	27.219
34	16.501	17.789	19.806	21.664	23.952	28.136
35	17.192	18.509	20.569	22.465	24.797	29.054
36	17.887	19.233	21.336	23.269	25.643	29.973
37	18.586	19.960	22.106	24.075	26.492	30.893
38	19.289	20.691	22.878	24.884	27.343	31.815
39	19.996	21.426	23.654	25.695	28.196	32.737
40	20.707	22.164	24.433	26.509	29.051	33.660
41	21.421	22.906	25.215	27.326	29.907	34.585
42	22.138	23.650	25.999	28.144	30.765	35.510
43	22.859	24.398	26.785	28.965	31.625	36.436
44	23.584	25.148	27.575	29.787	32.487	37.363
45	24.311	25.901	28.366	30.612	33.350	38.291

F-distribution

If X, Y are two independent χ^2−distributed random variables :

$$X \sim \chi^2(n)\ , \quad Y \sim \chi^2(m)$$

$$\longrightarrow \quad Z = \frac{X/n}{Y/m}$$

is called a F-distribution with first degree of freedom n, and 2nd degree of freedom m

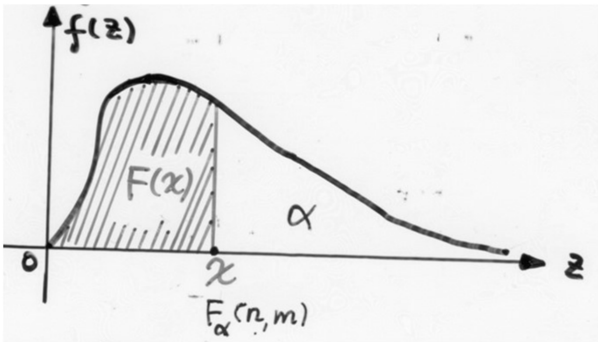
$$f(z) = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(\frac{n}{m}\right) \left(\frac{n}{m}z\right)^{\frac{n}{2}-1} \left(1+\frac{n}{m}z\right)^{-\frac{n+m}{2}} \qquad (z \geq 0)$$

$$E(z) = \frac{n}{m-2} \qquad (m > 2)$$

$$Var(z) = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)} \qquad (m > 4)$$

F-distribution

$$f(z) = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(\frac{n}{m}\right) \left(\frac{n}{m}z\right)^{\frac{n}{2}-1} \left(1+\frac{n}{m}z\right)^{-\frac{n+m}{2}} \qquad (z \geq 0)$$



F-distribution table

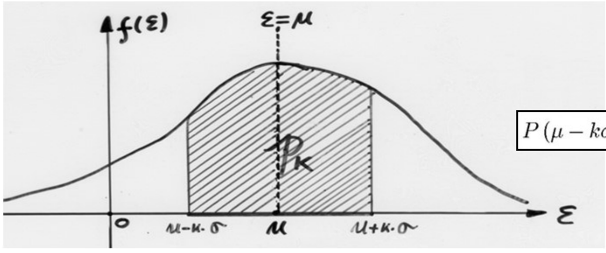
$P\{ F(n,m) > F_{\alpha}(n,m) \} = \alpha = 0.05$

m \ n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	161	200	216	225	230	234	237	239	241	242	243	244	245	245	246	246	247	247
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.70	8.69	8.68	8.67
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86	5.84	5.83	5.82
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62	4.60	4.59	4.58
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94	3.92	3.91	3.90
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51	3.49	3.48	3.47
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22	3.20	3.19	3.17
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97	2.96
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85	2.83	2.81	2.80
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72	2.70	2.69	2.67
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62	2.60	2.58	2.57
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53	2.51	2.50	2.48
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.46	2.44	2.43	2.41
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.45	2.42	2.40	2.38	2.37	2.35
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.35	2.33	2.32	2.30
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.35	2.33	2.31	2.29	2.27	2.26
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.31	2.29	2.27	2.25	2.23	2.22
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23	2.21	2.20	2.18
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.22	2.20	2.18	2.17	2.15
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.22	2.20	2.18	2.16	2.14	2.12
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.20	2.17	2.15	2.13	2.11	2.10
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.23	2.20	2.18	2.15	2.13	2.11	2.09	2.07
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.21	2.18	2.15	2.13	2.11	2.09	2.07	2.05
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.20	2.16	2.14	2.11	2.09	2.07	2.05	2.04
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.12	2.09	2.07	2.05	2.03	2.02
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.17	2.13	2.10	2.08	2.06	2.04	2.02	2.00
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12	2.09	2.06	2.04	2.02	2.00	1.99
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.08	2.05	2.03	2.01	1.99	1.97
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.06	2.04	2.01	1.99	1.98	1.96
32	4.15	3.29	2.90	2.67	2.51	2.40	2.31	2.24	2.19	2.14	2.10	2.07	2.04	2.01	1.99	1.97	1.95	1.94
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17	2.12	2.08	2.05	2.02	1.99	1.97	1.95	1.93	1.92
36	4.11	3.26	2.87	2.63	2.48	2.36	2.28	2.21	2.15	2.11	2.07	2.03	2.00	1.98	1.95	1.93	1.92	1.90
38	4.10	3.24	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02	1.99	1.96	1.94	1.92	1.90	1.88
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00	1.97	1.95	1.92	1.90	1.89	1.87

Confidence intervals

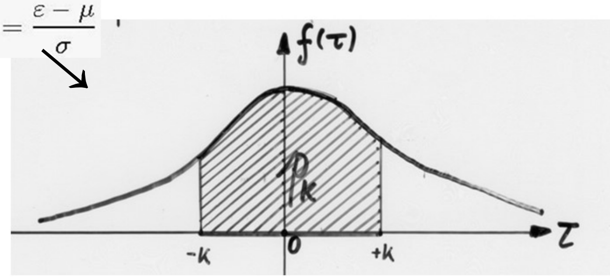
- A confidence interval of ε is an interval which contains the expectation μ
- The probability α that ε lies *outside* the confidence interval is called the *risk level*
- The probability that ε lies *inside* the confidence interval is 1-α , called the *confidence level*
- Confidence intervals are often taken as symmetrical intervals around the expectation μ

Symmetrical confidence intervals of $\varepsilon \sim N(\mu, \sigma^2)$



$$P\,(\mu - k\sigma \leq \varepsilon \leq \mu + k\sigma) = 1 - \alpha_k = p_k$$

$$\tau = \frac{\varepsilon - \mu}{\sigma}$$



$$p_k = P\,(-k \leq \tau \leq +k)$$

$$p_k = P\,(-k \leq \tau \leq +k) = 2\,[P\,(\tau \leq +k) - P\,(\tau \leq 0)] = 2\left[P\,(\tau \leq +k) - \frac{1}{2}\right] = 2 \cdot P\,(\tau \leq +k) - 1$$

Probability inside confidence intervals of $\varepsilon \sim N(\mu, \sigma^2)$

$$p_k = P\,(-k \leq \tau \leq +k) = 2\,[P\,(\tau \leq +k) - P\,(\tau \leq 0)] = 2\left[P\,(\tau \leq +k) - \frac{1}{2}\right] = 2 \cdot P\,(\tau \leq +k) - 1$$

$$p_1 = P\,(-1 \leq \tau \leq +1) = 2 \cdot 0.84134 - 1 = 68.27\%$$

$$p_2 = P\,(-2 \leq \tau \leq +2) = 2 \cdot 0.97725 - 1 = 95.45\%$$

$$p_3 = P\,(-3 \leq \tau \leq +3) = 2 \cdot 0.99865 - 1 = 99.73\%$$

Table 1.5: Confidence Intervals of Normal Distributions

k	*interval for* ε	*interval for* τ	*confidence level* ($p_k = 1 - \alpha_k$)	*risk level* (α_k)
1	$\mu \pm 1\sigma$	± 1	68.27 %	31.73 %
2	$\mu \pm 2\sigma$	± 2	95.45 %	4.55 %
3	$\mu \pm 3\sigma$	± 3	99.73 %	0.27 %
4	$\mu \pm 4\sigma$	± 4	99.99 %	0.01 %

$\longrightarrow$ 2σ or 3σ can be the maximum error tolerance !

Statistical Analysis (2/2)

- *Different distributions of random variables*
- *Confidence intervals*
- Hypothesis tests (statistical tests)
- Variance analysis
- Regression analysis (least squares fitting)

Sample mean and sample variance

Let $x_i \sim N(\mu, \sigma^2)$ $(i = 1, 2, \cdots, n)$ be n independent normally distributed variables with equal expectation μ and equal variance σ^2. A sample mean $\overline{x}$ and an estimated sample variance $\widehat{\sigma}^2$ can be calculated :

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \ , \quad \widehat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

$$\bar{x} \sim N(\mu, \frac{\sigma^2}{n}) \quad \longrightarrow \quad u = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

$$\frac{(n-1)\widehat{\sigma}^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{\overline{x} - x_i}{\sigma}\right)^2 \sim \chi^2(n-1) \longrightarrow t' = \frac{\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)\widehat{\sigma}^2}{\sigma^2} \cdot \frac{1}{n-1}}} = \frac{\overline{x} - \mu}{\widehat{\sigma}/\sqrt{n}} \sim t(n-1)$$

$$E\left(\widehat{\sigma}^2\right) = E\left\{\frac{\sigma^2}{n-1}\chi^2(n-1)\right\} = \sigma^2 \quad \longleftarrow$$

Unbiased estimate !

$$Var\left(\widehat{\sigma}^2\right) = E\left\{\left(\widehat{\sigma}^2 - \sigma^2\right)^2\right\} = \frac{2\sigma^4}{n-1}$$

2 measurement samples

$x_i \sim N(\mu_1, \sigma_1^2)\ (i = 1, 2, \cdots, n)$ $y_i \sim N(\mu_2, \sigma_2^2)\ (i = 1, 2, \cdots, m)$

$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\ ,\quad \widehat{\sigma}_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$ $\overline{y} = \frac{1}{m}\sum_{i=1}^{m} x_i\ ,\quad \widehat{\sigma}_2^2 = \frac{1}{m-1}\sum_{i=1}^{m}(y_i - \overline{y})^2$

$\frac{(n-1)\widehat{\sigma}_1^2}{\sigma_1^2} \sim \chi^2(n-1)$ $\frac{(m-1)\widehat{\sigma}_2^2}{\sigma_2^2} \sim \chi^2(m-1)$

$\downarrow$

$$F' = \frac{\frac{(n-1)\widehat{\sigma}_1^2}{\sigma_1^2} \cdot \frac{1}{n-1}}{\frac{(m-1)\widehat{\sigma}_2^2}{\sigma_2^2} \cdot \frac{1}{m-1}} = \frac{\sigma_2^2 \cdot \widehat{\sigma}_1^2}{\sigma_1^2 \cdot \widehat{\sigma}_2^2} \sim F(n-1, m-1)$$

Hypothesis tests (statistical tests)

- Make measurements / collect samples
- Compute sample statistics (mean, variance, etc)
- Define the null-hypothesis H_0 and the alternative hypothesis H_1
- Choose risk level α
- Find critical value from the distribution table
- Compare computed statistics against the critical value to decide whether to accept H_0 or H_1

Test expectation when variance is known

u-test

$x_1, x_2, \cdots, x_n$ are n independent measurements of the same normal distribution $N(\mu, \sigma^2)$

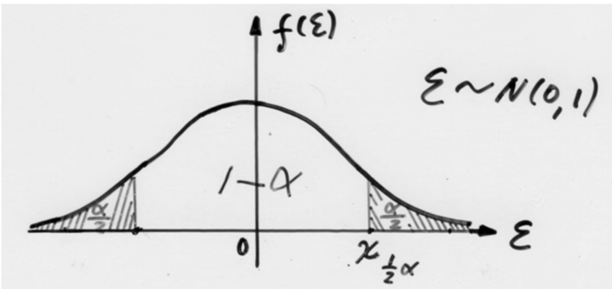
$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \bar{x} \sim N(\mu, \frac{\sigma^2}{n}) \quad \text{and} \quad \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Null hypothesis H_0 :	$\overline{x} = \mu$
Alternative hypothesis H_1 :	$\overline{x} \neq \mu$

Choose risk level $\alpha = 5\%$

Test expectation when variance is known

Find the cirical value $x_{\frac{1}{2}\alpha}$ of $N(0,1)$



$$P\left\{\left|\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}\right| \leq x_{\frac{1}{2}\alpha}\right\} = 1-\alpha \quad or: P\{|\bar{x}-\mu| \leq c_1\} = 1-\alpha$$

Compare critical value $x_{\frac{1}{2}\alpha}$ with computed $\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$

Decide whether to accept H_0 or H_1

Test expectation when variance is unknown

t-test

$x_1, x_2, \cdots, x_n$ are n independent measurements of the same normal distribution $N(\mu, \sigma^2)$

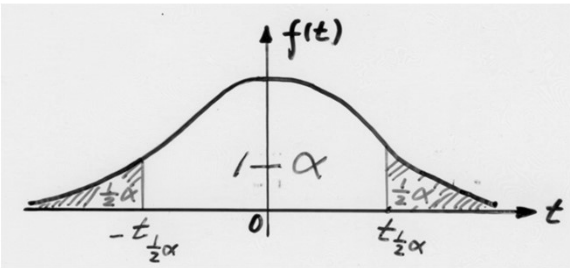
$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \widehat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \qquad \frac{\overline{x}-\mu}{\widehat{\sigma}/\sqrt{n}} \sim t(n-1)$$

Null hypothesis H_0 :	$\overline{x} = \mu$
Alternative hypothesis H_1 :	$\overline{x} \neq \mu$

Choose risk level α =5%

Test expectation when variance is unknown

Find the cirical value $t_{\frac{1}{2}\alpha}(n-1)$ of t-distribution



$$P\left\{\frac{\overline{x}-\mu}{\widehat{\sigma}/\sqrt{n}} \leq t_{\frac{1}{2}\alpha}(n-1)\right\} = 1-\alpha \quad or:\ P\{|\bar{x}-\mu| \leq c_2\} = 1-\alpha$$

Compare $\frac{\overline{x}-\mu}{\widehat{\sigma}/\sqrt{n}}$ with $t_{\frac{1}{2}\alpha}(n-1)$

Decide whether to accept H_0 or H_1

Test whether two samples have the same variance

F-test

$x_i \sim N(\mu_1, \sigma_1^2)\ (i = 1, 2, \cdots, n)$ $\qquad$ $y_i \sim N(\mu_2, \sigma_2^2)\ (i = 1, 2, \cdots, m)$

$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\ ,\quad \widehat{\sigma}_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$ $\qquad$ $\overline{y} = \frac{1}{m}\sum_{i=1}^{m} x_i\ ,\quad \widehat{\sigma}_2^2 = \frac{1}{m-1}\sum_{i=1}^{m}(y_i - \overline{y})^2$

$\frac{(n-1)\widehat{\sigma}_1^2}{\sigma_1^2} \sim \chi^2(n-1)$ $\qquad$ $\frac{(m-1)\widehat{\sigma}_2^2}{\sigma_2^2} \sim \chi^2(m-1)$

$\downarrow$

$$F' = \frac{\frac{(n-1)\widehat{\sigma}_1^2}{\sigma_1^2} \cdot \frac{1}{n-1}}{\frac{(m-1)\widehat{\sigma}_2^2}{\sigma_2^2} \cdot \frac{1}{m-1}} = \frac{\sigma_2^2 \cdot \widehat{\sigma}_1^2}{\sigma_1^2 \cdot \widehat{\sigma}_2^2} \sim F(n-1, m-1)$$

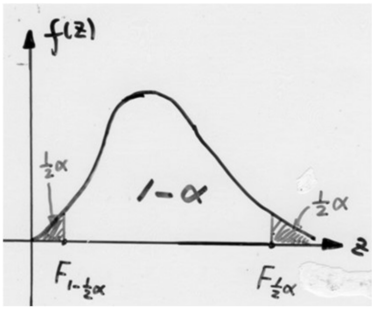
? $\boxed{H_0:\ \sigma_1^2 = \sigma_2^2\ ; \quad H_1: \sigma_1^2 \neq \sigma_2^2}$

Test whether two samples have the same variance

Under H0, $\qquad \frac{\widehat{\sigma}_1^2}{\widehat{\sigma}_2^2} \sim F(n-1, m-1)$

Choose risk level α

$$P\left\{F_{1-\frac{1}{2}\alpha}(n-1, m-1) < \frac{\widehat{\sigma}_1^2}{\widehat{\sigma}_2^2} < F_{\frac{1}{2}\alpha}(n-1, m-1)\right\} = 1 - \alpha$$



Order so that $\frac{\widehat{\sigma}_1^2}{\widehat{\sigma}_2^2} > 1$

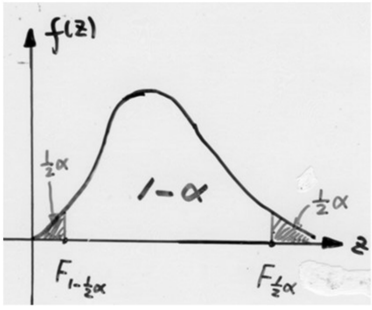
$F_{1-\frac{1}{2}\alpha}(n-1, m-1)$ is always smaller than 1 for small α.

$$P\left\{\frac{\widehat{\sigma}_1^2}{\widehat{\sigma}_2^2} < F_{\frac{1}{2}\alpha}(n-1, m-1)\right\} = 1 - \alpha$$

Test whether two samples have the same variance

Example. n=41, m=61, α=5%

$$F_{\frac{1}{2}\alpha}(n-1,m-1) = 1.74\ , \quad \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = (4.2/3.0)^2 = 1.96 > F_{\frac{1}{2}\alpha}(n-1,m-1)$$



Conclusion:
the two samples have *significantly different* variances at risk level α=5%

Variance Analysis

Factors (groups)	Measurements	No.	Distributions
1	$l_{11}\ l_{12}\ l_{13}\ \cdots\ l_{1n_1}$	n_1	$N(\mu_1, \sigma^2)$
2	$l_{21}\ l_{22}\ l_{23}\ \cdots\ l_{2n_2}$	n_2	$N(\mu_2, \sigma^2)$
$\vdots$			
i	$l_{i1}\ l_{i2}\ l_{i3}\ \cdots\ l_{in_i}$	n_i	$N(\mu_i, \sigma^2)$
$\vdots$			
m	$l_{m1}\ l_{m2}\ l_{m3}\ \cdots\ l_{mn_m}$	n_m	$N(\mu_m, \sigma^2)$

Hypothesis: $H_0 : \mu_1 = \mu_2 = \cdots = \mu_m$?

Internal variation

- Within each group, compute mean and variance

$$\widehat{\mu}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} \ell_{ij} \quad (i = 1, 2, 3, \cdots, m) \qquad \widehat{\sigma}_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i} (\ell_{ij} - \widehat{\mu}_i)^2 \quad (i = 1, 2, 3, \cdots, m)$$

- Internal variation

$$S_I^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i} (\ell_{ij} - \widehat{\mu}_i)^2 = \sum_{i=1}^{m} \{(n_i - 1) \cdot \widehat{\sigma}_i^2\}$$

- Number of measurements: $n = \sum_{i=1}^{m} n_i = n_1 + n_2 + \cdots + n_m$
- Number of group averages: m
- Internal variance: $\widehat{\sigma}_I^2 = \frac{S_I^2}{n - m} = \frac{1}{n - m}\sum_{i=1}^{m}\sum_{j=1}^{n_i} (\ell_{ij} - \widehat{\mu}_i)^2$

$$\frac{S_I^2}{\sigma^2} \sim \chi^2(n - m)$$

External variation

- Overall mean $\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{n_i} \ell_{ij} = \frac{1}{n}\sum_{i=1}^{m} (n_i \cdot \widehat{\mu}_i)$
- External variation $S_E^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i} (\widehat{\mu}_i - \widehat{\mu})^2 = \sum_{i=1}^{m} n_i (\widehat{\mu}_i - \widehat{\mu})^2$

$$\frac{S_E^2}{\sigma^2} \sim \chi^2(m - 1)$$

- External variance $\widehat{\sigma}_E^2 = \frac{1}{m - 1}\sum_{i=1}^{m} \left\{ n_i (\widehat{\mu}_i - \widehat{\mu})^2 \right\} = \frac{S_E^2}{m - 1}$

Total variation

- Total variation

$$S_T^2 = S_T^2(\widehat{\mu}) = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(\ell_{ij}-\widehat{\mu})^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(\ell_{ij}-\widehat{\mu}_i)^2 + \sum_{i=1}^{m} n_i\,(\widehat{\mu}_i-\widehat{\mu})^2$$

$$S_T^2 = S_I^2 + S_E^2$$

$$\frac{S_T^2}{\sigma^2} = \frac{S_I^2}{\sigma^2} + \frac{S_E^2}{\sigma^2} \sim \chi^2(n-1)$$

- Total variance

$$\widehat{\sigma}_T^2 = \frac{S_T^2}{n-1} = \frac{S_I^2+S_E^2}{n-1} = \frac{(n-m)\cdot\widehat{\sigma}_I^2 + (m-1)\cdot\widehat{\sigma}_E^2}{(n-m)+(m-1)}$$

Summary of variations

Type	Variations	Degree of Freedom	Distributions	Variances
Internal	$S_I^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(\ell_{ij}-\hat{\mu}_i)^2$	$n-m$	$\frac{S_I^2}{\sigma^2} \sim \chi^2(n-m)$	$\hat{\sigma}_I^2 = \frac{S_I^2}{n-m}$
External	$S_E^2 = \sum_{i=1}^{m} n_i(\hat{\mu}_i-\hat{\mu})^2$	$m-1$	$\frac{S_E^2}{\sigma^2} \sim \chi^2(m-1)$	$\hat{\sigma}_E^2 = \frac{S_E^2}{m-1}$
Total	$S_T^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(\ell_{ij}-\hat{\mu})^2$ $= S_I^2 + S_E^2$	$n-1$ $= (n-m)+(m-1)$	$\frac{S_T^2}{\sigma^2} \sim \chi^2(n-1)$ $= \frac{S_I^2}{\sigma^2} + \frac{S_E^2}{\sigma^2}$	$\hat{\sigma}_T^2 = \frac{S_T^2}{n-1}$

$$F_0 = \frac{\frac{S_E^2}{m-1}}{\frac{S_I^2}{n-m}} = \frac{\hat{\sigma}_E^2}{\hat{\sigma}_I^2} \sim F(m-1,\ n-m)$$

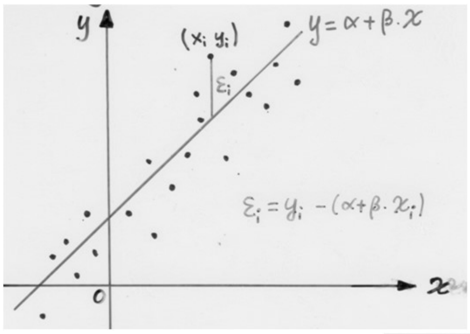
Variance analysis

- Comparison of the external and the internal variances:

$$F_0 = \frac{S_E^2/(m-1)}{S_I^2/(n-m)} = \frac{\widehat{\sigma}_E^2}{\widehat{\sigma}_I^2} \sim F(m-1, n-m)$$

- For chosen risk level α, find the critical value
- If $F_0 > F_\alpha(m-1, n-m)$, reject H_0 at risk level α

Regression analysis

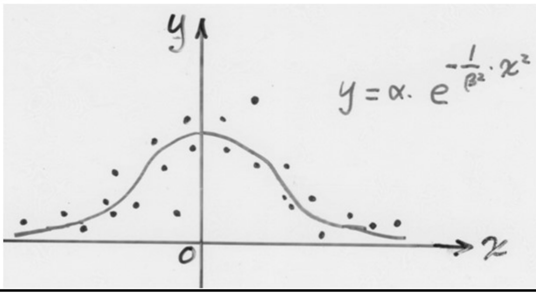


To fit a mathematical function to random measurements.

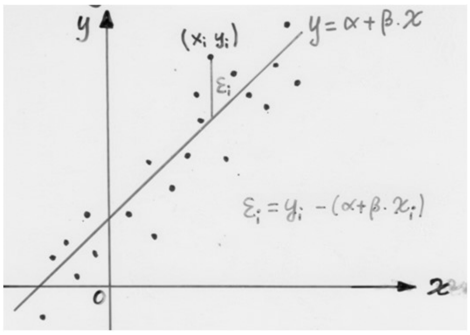
To find systematic trend from irregular data.

With the help of the least squares principle:

$$\sum_{i=1}^{n} \varepsilon_i^2 = \text{minimum}$$



Linear regression of one variable



$$y_i = \alpha + \beta \cdot x_i + \varepsilon_i$$

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \alpha - \beta \cdot x_i)^2 = \text{minimum}$$

$$\frac{\partial}{\partial \alpha} \left\{ \sum_{i=1}^{n} (y_i - \alpha - \beta \cdot x_i)^2 \right\}_{\alpha=\widehat{\alpha},\ \beta=\widehat{\beta}} = \sum_{i=1}^{n} 2(y_i - \widehat{\alpha} - \widehat{\beta} \cdot x_i)\,(-1) = 0$$

$$\frac{\partial}{\partial \beta} \left\{ \sum_{i=1}^{n} (y_i - \alpha - \beta \cdot x_i)^2 \right\}_{\alpha=\widehat{\alpha},\ \beta=\widehat{\beta}} = \sum_{i=1}^{n} 2(y_i - \widehat{\alpha} - \widehat{\beta} \cdot x_i)\,(-x_i) = 0$$

Computation of regression coefficients

$$\begin{bmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{bmatrix} \begin{bmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

$$q_{11} = n$$
$$q_{12} = \sum_{i=1}^{n} x_i$$
$$q_{22} = \sum_{i=1}^{n} x_i^2$$
$$w_1 = \sum_{i=1}^{n} y_i$$
$$w_2 = \sum_{i=1}^{n} x_i y_i$$

$$\begin{bmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{bmatrix} = \frac{1}{q_{11}q_{22} - q_{12}^2} \begin{bmatrix} q_{22} & -q_{12} \\ -q_{12} & q_{11} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \frac{1}{q_{11}q_{22} - q_{12}^2} \begin{bmatrix} q_{22}w_1 - q_{12}w_2 \\ -q_{12}w_1 + q_{11}w_2 \end{bmatrix}$$

$$\widehat{\varepsilon}_i = y_i - \widehat{\alpha} - \widehat{\beta} \cdot x_i \qquad \widehat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} \widehat{\varepsilon}_i^2 = \frac{1}{n-2} \sum_{i=1}^{n} \left(y_i - \widehat{\alpha} - \widehat{\beta} \cdot x_i \right)^2$$

$$\sigma_{\widehat{\alpha}} = \widehat{\sigma} \cdot \sqrt{\frac{q_{22}}{q_{11}q_{22} - q_{12}^2}}, \qquad \sigma_{\widehat{\beta}} = \widehat{\sigma} \cdot \sqrt{\frac{q_{11}}{q_{11}q_{22} - q_{12}^2}}$$

$$\frac{\widehat{\alpha}}{\sigma_{\widehat{\alpha}}} \sim t(n-2), \qquad \frac{\widehat{\beta}}{\sigma_{\widehat{\beta}}} \sim t(n-2) \longrightarrow$$

Are the computed coefficients *statistically significant* ?

Standardfel bör vara minst 3x större än beräknat fel

Reduce non-linear relations into linear regression model

$$y_i = \frac{x_i}{\alpha \cdot x_i + \beta} \longrightarrow Y_i = \frac{1}{y_i}, \quad X_i = \frac{1}{x_i} \longrightarrow Y_i = \alpha + \beta \cdot X_i$$

$$y_i = \alpha \cdot e^{-\frac{x_i^2}{\beta^2}} \longrightarrow \ln y_i = \ln \alpha - \frac{1}{\beta^2} \cdot x_i^2 \longrightarrow Y_i = a + b \cdot X_i$$

$$X_i = x_i^2, \quad Y_i = \ln y_i, \quad a = \ln \alpha, \quad b = -\frac{1}{\beta^2}$$

$$\widehat{\alpha} = e^{\widehat{a}}, \quad \widehat{\beta} = \sqrt{-\frac{1}{\widehat{b}}}$$

$$\sigma_i^2 = a^2 + b^2 \cdot s_i^2$$

$$y_i = \sigma_i^2, \quad x_i = s_i^2, \quad \alpha = a^2, \quad \beta = b^2$$

$$y_i = \alpha + \beta \cdot x_i$$

Constant and scale errors of EDM instruments

$$\sigma_i^2 = a^2 + b^2 \cdot s_i^2$$

i	s_i (km)	σ_i (cm)
1	0.5	2.9
2	1.2	3.1
3	1.9	3.2
4	3.0	3.4
5	3.7	3.5
6	4.4	3.7
7	4.9	3.8
8	5.1	4.1
9	5.7	4.2
10	6.0	4.4

Constant and scale errors of EDM instruments

$$y_i = \sigma_i^2, \quad x_i = s_i^2, \quad \alpha = a^2, \quad \beta = b^2$$

$$y_i = \alpha + \beta \cdot x_i$$

i	$x_i = s_i^2\ (km^2)$	$y_i = \sigma_i^2\ (cm^2)$	x_i^2	$x_i \cdot y_i$
1	0.25	8.41	0.0625	2.1025
2	1.44	9.61	2.0736	13.8384
3	3.61	10.24	13.0321	36.9664
4	9.00	11.56	81.0000	104.0400
5	13.69	12.25	187.4161	167.7025
6	19.36	13.69	374.8096	265.0384
7	24.01	14.44	576.4801	346.7044
8	26.01	16.81	676.5201	437.2281
9	32.49	17.64	1055.6001	573.1236
10	36.00	19.36	1296.000	696.9600
$\sum$	165.86	134.01	4262.9942	2643.7043

Constant and scale errors of EDM instruments

$$\begin{bmatrix} 10 & 165.86 \\ 165.86 & 4262.9942 \end{bmatrix} \begin{bmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{bmatrix} = \begin{bmatrix} 134.01 \\ 2643.7043 \end{bmatrix}$$

$$\begin{bmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{bmatrix} = \frac{1}{15120.403} \begin{bmatrix} 4262.9942 & -165.86 \\ -165.86 & 10 \end{bmatrix} \begin{bmatrix} 134.01 \\ 2643.7043 \end{bmatrix} = \begin{bmatrix} 8.783 \\ 0.278 \end{bmatrix}$$

$$\widehat{\sigma}^2 = \frac{1}{10-2} \sum_{i=1}^{n} \left(y_i - \widehat{\alpha} - \widehat{\beta} \cdot x_i \right)^2 = 0.3782 \quad \rightarrow \quad \widehat{\sigma} = \pm 0.615\ (cm^2)$$

$$\sigma_{\widehat{\alpha}} = \widehat{\sigma} \cdot \sqrt{\frac{4262.9942}{15120.403}} = 0.326\ (cm^2), \quad \sigma_{\widehat{\beta}} = \widehat{\sigma} \cdot \sqrt{\frac{10}{15120.403}} = 0.015\ \ (cm/km)^2$$

$$\widehat{a} = \sqrt{\widehat{\alpha}} = \sqrt{8.783} = 2.96\ cm, \quad \widehat{b} = \sqrt{\widehat{\beta}} = \sqrt{0.278} = 0.53\ (cm/km) = 5.3\ ppm$$