فصل دوم: آشنایی با روش‌های آماری در تحلیل داده‌های اندازه‌گیری شده

روش‌ها و مطالب فصل:
- محاسبه میانگین و انحراف معیار
- تابع چگالی احتمال (PDF)
- محاسبه احتمال
- آمار در مجموعه داده‌های محدود (finite sized)
- توزیع Student's t
- برازش منحنی / رگرسیون

* خطای اندازه‌گیری: اختلاف بین مقدار اندازه‌گیری شده و مقدار واقعی متغیر مورد نظر

$$error = x_{measured} - \underbrace{x_{actual}}_{\text{true value}}$$

true value مقدار واقعی

* انواع خطا و عدم قطعیت در سیستم‌های اندازه‌گیری:

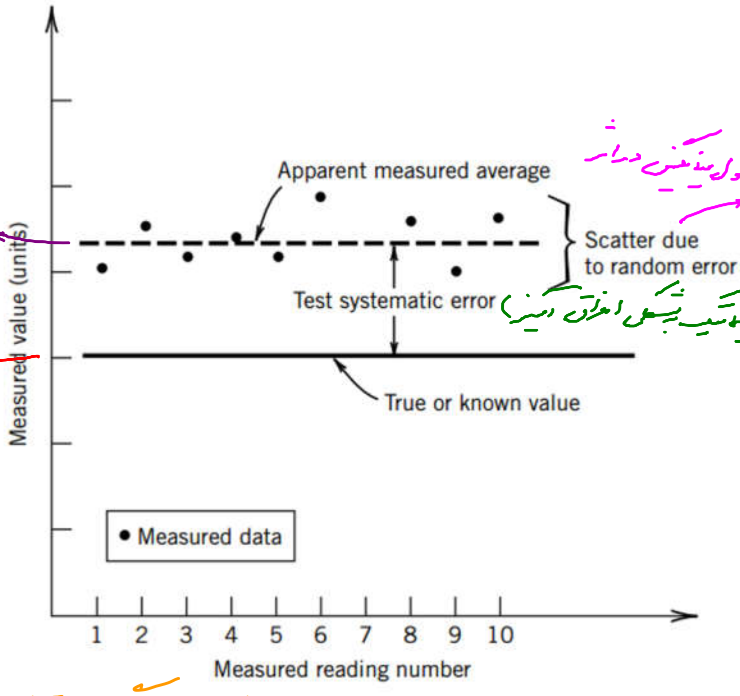
الف) خطای سیستماتیک یا معین (Deterministic): خطاهایی که موجب ایجاد تفاوت مقدار در میانگین نتایج تعداد زیادی اندازه‌گیری مستقل از مقدار واقعی کمیت مورد نظر می‌شود.

(به عنوان مثال خطاهای ایجاد شده در نتیجه به هم خوردن کالیبراسیون، طراحی نامناسب سنسور، اثرات بایاس، خطاهای ایجاد شده در اثر ورودی‌های تداخلی (تغییر دهنده) و ...)

ب) خطای تصادفی (Random) : خطاهایی که سبب می‌شود که مقادیر اندازه‌گیری شده در شرایط یکسان تکرارپذیر نباشند.

← خطای تصادفی با مفاهیم آماری نظیر میانگین، انحراف معیار، بازه عدم قطعیت و سطح اطمینان مشخص می‌شود.

← معمولاً توسط شرکت سازنده به صورت مشخصه نویز سنسور/سیستم اندازه‌گیری ارائه می‌شود.



میانگین قرائت‌های مستقل انجام شده

مقدار واقعی کمیت

پراکندگی داده‌ها حول میانگین در اثر خطای تصادفی

(خطای سیستماتیک به شکل انحراف اریب)

دفعات مختلف اندازه‌گیری و قرائت

* شاخص های آماری :

- به نتایج خروجی از اندازه گیری های صورت گرفته توسط یک دستگاه اندازه گیری مشاهده (observation) $x_1, x_2, \dots, x_n$

و به مجموعه یک سری مشاهده، نمونه (sample) گویند.

- میانگین را مقدار متوسط نمونه :

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^{n} x_i$$

← میانگین نمونه، مقدار قابل انتظار مشاهده بعدی است.

- انحراف از معیار نمونه (samples standard deviation) :

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

← معیاری از پراکندگی داده های یک نمونه است.

← زمانی که تعداد قرائت سنسور (مشاهدات) (n) بزرگ شده و به بی نهایت میل کند،

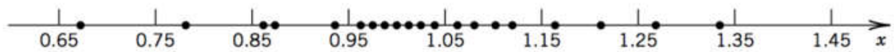
σ_n : معیاری از عدم قطعیت یا دقت سنسور است.

- عموماً می‌توان از رابطه زیر برای تخمین بازه مقدار اندازه‌گیری شده در مشاهده‌ها استفاده کرد:

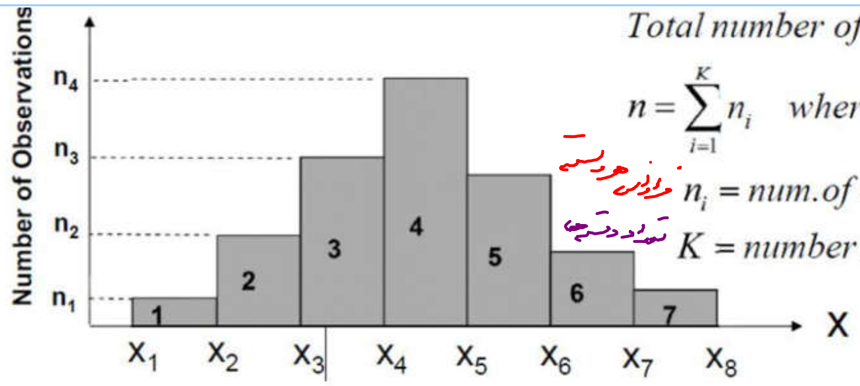
(با درجه/ضریب اطمینان P٪) بازه عدم قطعیت ± متوسط نمونه = مقدار اندازه‌گیری/قرائت شده ($x_{measured}$)

که در آن، درجه اطمینان (confidence level) عبارت است از احتمالی اینکه یک مقدار قرائت شده در محدوده مشخصی از خطا قرار گیرد.

- عموماً کمیت اندازه‌گیری شده حول مقدار متوسط آن دارای خطاست. این خطا دارای یک توزیع است که با تابع چگالی احتمال (probability density function) مشخص می‌شود.



- می‌توان از هیستوگرام جهت مشاهده توزیع احتمال یک پدیده استفاده کرد.



Total number of observations:

$$n = \sum_{i=1}^{K} n_i \quad \text{where:}$$

فراوانی هر دسته n_i = num. of obs. in each bin

تعداد دسته‌ها K = number of bins (7)

■ The "best" histogram is obtained using ***Sturgis' rule***

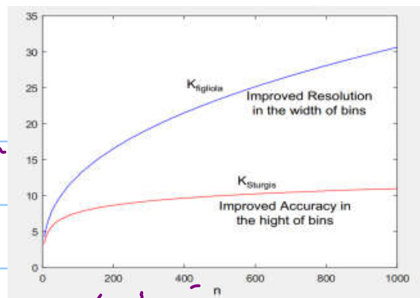
$$K = 1 + 3.33\log_{10} n$$ ← تخمین تعداد دسته‌ها از روش استورجس

■ An alternative ***rule*** *often used (in the Figliola-Beasley) is*

$$K = 1 + 1.87(n-1)^{0.4}$$ ← پیشنهاد داده شده در کتاب معمولا

K : number of bins

n : total number of observations.



K تعداد دسته‌ها

تعداد داده‌ها

- مثال: متغیر تصادفی x ، ۲۰ مرتبه اندازه‌گیری شده است. هیستوگرام و توزیع فراوانی آن را رسم کنید.

i	x_i	i	x_i
1	0.98	11	1.02
2	1.07	12	1.26
3	0.86	13	1.08
4	1.16	14	1.02
5	0.96	15	0.94
6	0.68	16	1.11
7	1.34	17	0.99
8	1.04	18	0.78
9	1.21	19	1.06
10	0.86	20	0.96

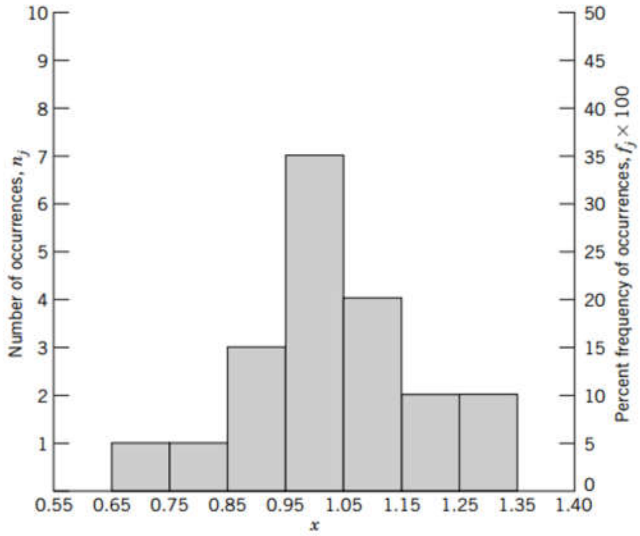
$N = 20$

$K = 1.87(N-1)^{0.4} + 1 = 7$ — تعداد دسته‌ها

Fig.

فراوانی نسبی هر دسته

j	Interval	n_j	$f_j = n_j/N$
1	$0.65 \le x_i < 0.75$	1	0.05
2	$0.75 \le x_i < 0.85$	1	0.05
3	$0.85 \le x_i < 0.95$	3	0.15
4	$0.95 \le x_i < 1.05$	7	0.35
5	$1.05 \le x_i < 1.15$	4	0.20
6	$1.15 \le x_i < 1.25$	2	0.10
7	$1.25 \le x_i < 1.35$	2	0.10



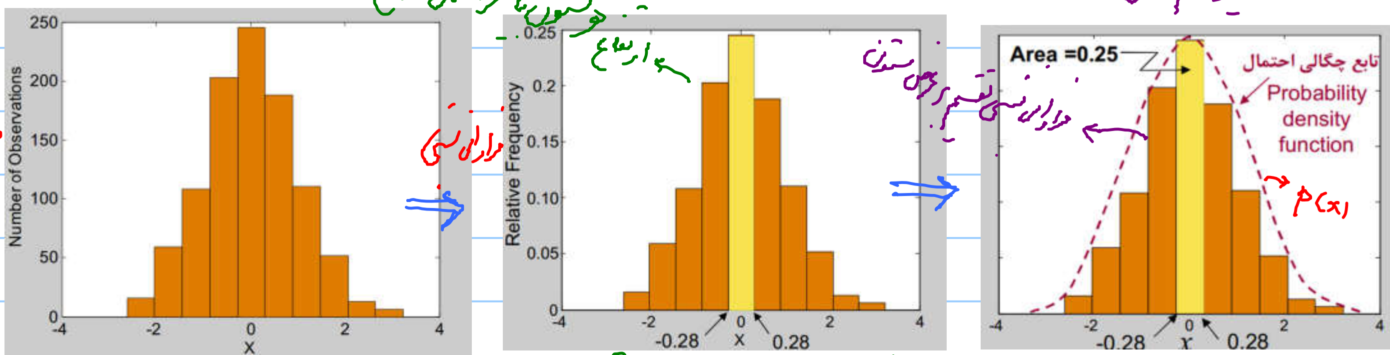
درصد فراوانی نسبی هر دسته

- اگر تعداد مشاهدات و تعداد دسته‌ها به سمت بی‌نهایت میل کند، می‌توان تابع توزیع چگالی احتمال (PDF) را اینگونه استخراج نمود:

$$(\text{با حروف کوچک})\quad p(x) = \lim_{\delta x \to 0} \frac{\overbrace{n_j}^{\text{فراوانی مطلق هر دسته}}}{\underbrace{N}_{\text{تعداد کل داده‌ها}}\underbrace{(2\delta x)}_{\text{عرض هر دسته}}}$$

ستون‌ها برابر احتمال وقوع x در آن بازه است.

هیستوگرام نرمال شده:



فراوانی مطلق

فراوانی نسبی

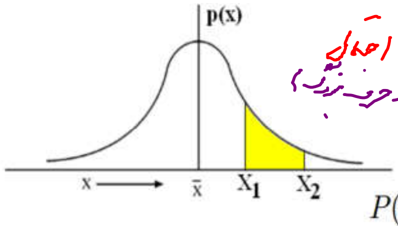
ارتفاع

فراوانی نسبی تقسیم بر عرض ستون

$p(x)$

$P(-0.28 < x < 0.28) \simeq 25\%$

- سطح زیر منحنی تابع چگالی احتمال از x_1 تا x_2، برابر احتمال رخ دادن x در محدوده (x_1, x_2) است:



احتمال (حروف بزرگ) $P(X_1 \le X \le X_2) = \int_{X_1}^{X_2} p(x)dx$ تابع چگالی احتمال

$$P(-\infty < X < +\infty) = \int_{-\infty}^{+\infty} p(x)dx = 1$$

- صرف‌نظر از رابطه $p(x)$، داریم:

$$\mu = x' = \int_{-\infty}^{\infty} xp(x)dx \qquad \sigma^2 = \int_{-\infty}^{\infty} (x-x')^2 p(x)dx$$

مقدار متوسط واقعی True mean value — True variance

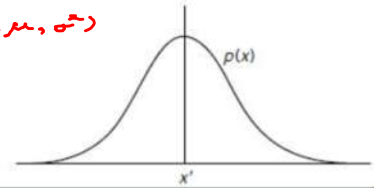
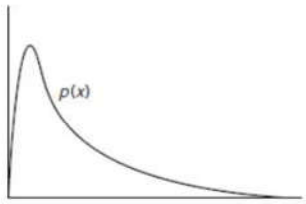
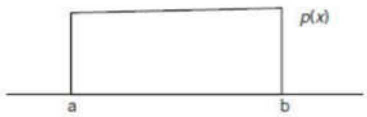
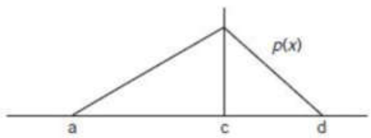
که (در مسائل اندازه‌گیری به آن خروجی واقعی (true output) هم گفته می‌شود

$(n \to \infty : \bar{x}_n \to \mu)$

* نمونه‌ای از توابع توزیع احتمال پرکاربرد

توزیع احتمال نرمال (گوسی)

پرکاربردترین

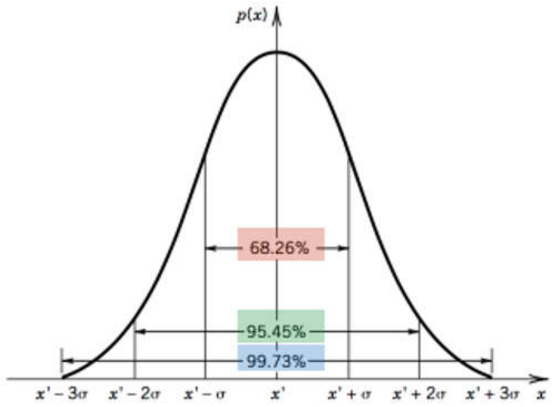
Distribution	Applications	Density Function	Shape
Normal	Most physical properties that are continuous or regular in time or space with variations due to random error	$p(x) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\left[-\frac{1}{2}\frac{(x-x')^2}{\sigma^2}\right]$ = مقدار متوسط √واریانس — میانگین (واقعی)	
Log normal	Failure or durability projections; events whose outcomes tend to be skewed toward the extremity of the distribution	$p(x) = \frac{1}{x\sigma(2\pi)^{1/2}} \exp\left[-\frac{1}{2}\frac{\ln(x-x')^2}{\sigma^2}\right]$	
Rectangular	Processes in which a likely outcome falls in the range between minimum value a and maximum value b occurring with equal probability	$p(x) = \frac{1}{b-a}$ where $a \le x \le b$, otherwise $p(x) = 0$	
Triangular	Process in which the likely outcome x falls between known lower bound a and upper bound b, with a peak value or mode c; used when population information is sparse	$p(x) = \frac{2(x-a)}{(b-a)(c-a)}$ for $a \le x \le c$ $= \frac{2(b-x)}{(b-a)(b-c)}$ for $c < x \le b$	

- توزیع احتمال بسیاری از پدیده‌ها (در صنعت) و اتفاقات، از جمله اندازه‌گیری و خروجی سنسورها نرمال است و با

نرمال مدل می‌شود.

- توزیع احتمال نرمال (گوسی) فقط به مقدار متوسط و انحراف از معیار وابسته است.

← انحراف از معیار کوچکتر، باعث (دقت بیشتر شدن سیستم اندازه‌گیری) است.



$$P(\bar{x}-\sigma \le X \le \bar{x}+\sigma) = \int_{\bar{x}-\sigma}^{\bar{x}+\sigma} p(x)dx = 0.68 = \int_{-1}^{1} p(z)\,dz$$

$$P(\bar{x}-2\sigma \le X \le \bar{x}+2\sigma) = \int_{\bar{x}-2\sigma}^{\bar{x}+2\sigma} p(x)dx = 0.95 = \int_{-2}^{2} p(z)\,dz$$

$$P(\bar{x}-3\sigma \le X \le \bar{x}+3\sigma) = \int_{\bar{x}-3\sigma}^{\bar{x}+3\sigma} p(x)dx = 0.99 = \int_{-3}^{3} p(z)\,dz$$

- اگر $x' = \mu = 0$ و $\sigma = 1$ باشد، توزیع را توزیع نرمال استاندارد گویند.

- هر توزیع نرمالی را می‌توان با تغییر متغیر، تبدیل به یک توزیع نرمال استاندارد کرد.

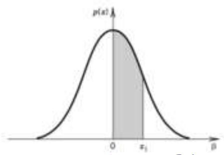
$$\boxed{z = \frac{x - x'}{\sigma}} \Rightarrow p(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

Probability Values for Normal Error Function

z ← رقم دوم اعشار

رقم اول و اعشار z

$z_1 = \frac{x_1 - x'}{\sigma}$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1809	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2794	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4292	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4758	0.4761	0.4767
2.0	0.4772	0.4778	0.4803	0.4788	0.4793	0.4799	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.49865	0.4987	0.4987	0.4988	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990



متوسط — عدم قطعیت — درجه اطمینان

$$x_i = x' \pm z_1\sigma \quad (P\%)$$

← احتمال استخراج شده از جدول که حاصل شده از انتگرال $\left(\int_a^b p\,dx\right)$

$x_{\text{measured}} = \mu \pm \sigma$ (with a confidence level of 68%)

$x_{\text{measured}} = \mu \pm 2\sigma$ (with a confidence level of 95%)

$x_{\text{measured}} = \mu \pm 3\sigma$ (with a confidence level of 99.7%)

مثال: اگر میانگین و انحراف معیار یک سیگنال ولتاژ به ترتیب $\bar{x} = 8.5\ V$ و $\sigma = 1.5\ V$ باشد، با چه احتمالی

در یک اندازه‌گیری مقدار ولتاژ قرائت شده بین 10.0 و 11.5 V است؟

$$P(10.0 \le x \le 11.5) = P(8.5 \le x \le 11.5) - P(8.5 \le x \le 10.0) = 0.4772 - 0.3413 = 14\%$$

$$z = 0 \qquad z_2 = \frac{11.5 - 8.5}{1.5} = 2 \qquad z_1 = \frac{10.0 - 8.5}{1.5} = 1.0$$

$$P(0 \le z \le 2) = 0.4772 \qquad P(0 \le z \le 1) = 0.3413$$

مثال- یک دستگاه تست کشش، نیروی مورد نیاز جهت پاره کردن یک نمونه ده‌تایی از کابل‌های فولادی

را مطابق روبرو اندازه‌گیری نموده است.

$x_1 = 4.3\ KN,\ x_2 = 4.5\ KN,\ x_3 = 4.7\ KN,\ x_4 = 4.2\ KN,\ x_5 = 4.5\ KN$

$x_6 = 4.6\ KN,\ x_7 = 4.4\ KN,\ x_8 = 4.6\ KN,\ x_9 = 4.9\ KN,\ x_{10} = 4.5\ KN$

الف) میانگین و انحراف معیار نمونه را محاسبه نمایید.

ب) با چه ضریب اطمینانی، می‌توان عدم قطعیت اندازه‌گیری را 0.2 KN اعلام نمود؟ (≡ احتمال اینکه

یک اندازه‌گیری در فاصله 0.2 KN نسبت به مقدار متوسط قرار گیرد چقدر است؟)

پ) عدم قطعیت اندازه‌گیری با ضریب اطمینان 95% چقدر است؟ (≡ 95% از اندازه‌گیری‌ها در

چه محدوده‌ای از میانگین قرار خواهد داشت؟)

الف) میانگین نیروی نمونه‌ها: $\bar{x}_n = \dfrac{\sum_{i=1}^{10} x_i}{10} = \underline{4.5}$ kN $= \mu$ (فرض می‌کنیم میانگین واقعی هم همین مقدار است!)

$$\sigma_n = \sqrt{\frac{\sum_{i=1}^{10}(x_i - \bar{x}_n)^2}{10}} = \underline{0.19}\ \text{kN}$$

میزان متوسط یا دقت در حد نیروی کشش نمونه‌ها با درجه اطمینان ۶۸٪ (با فرض نرمال بودن نمونه‌ها عبارت است از: $x = 4.5 \pm 0.19$

ب)

$$x = \mu \pm \underbrace{Z \cdot \sigma}_{\text{عدم قطعیت}}$$

$$\rightarrow Z = \frac{x-\mu}{\sigma} = \frac{0.2}{0.19} \Rightarrow Z = 1.06$$

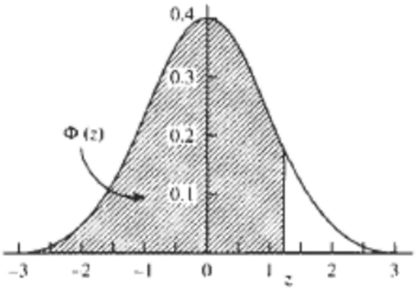
$$P(-1.06 < x < 1.06) = 2\int_0^{1.06} p(z)\,dz \overset{\text{جدول}}{=\!=} 2 \times 0.355 = 0.71 = \underline{71\%}$$

ج)

$$P(-Z < x < \underbrace{Z}_{\text{از جدول}}) = 0.98 \Rightarrow \int_0^{Z} p(z)\,dz = \frac{0.98}{2} = 0.49$$

$$\overset{\text{جدول}}{\Longrightarrow} Z = \frac{x-\mu}{\sigma_n} = 2.33 \Rightarrow \underbrace{\text{عدم قطعیت}}_{\text{precision}} = \pm 2.33 \times \sigma_n = \pm 0.44\ \text{kN}$$

* تابع توزیع تجمعی نرمال :



$$\Phi(\psi) = P(z < \psi) = \int_{-\infty}^{\psi} p(z)dz = \int_{-\infty}^{\psi} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

$$P(\alpha < z < \beta) = \Phi(\beta) - \Phi(\alpha), \qquad \Phi(-z) = 1 - \Phi(z)$$

رقم دوم اعشار

یکان و رقم اول اعشار

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5399	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	0.8364	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	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	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	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	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	0.9875	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
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
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	0.9078	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9085	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
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

* آمار در مجموعه داده های محدود (Statistics of Finite-Sized Datasets)

- جهت دستیابی به مقدار واقعی میانگین و انحراف معیار یک متغیر تصادفی، می بایست تعداد داده های بسیار بسیار زیادی ($n \to \infty$) داشته باشیم. در این حالت میانگین و انحراف معیار محاسبه شده، صرفاً تخمینی از مقادیر واقعی هستند.

میانگین نمونه $\bar{x}_n = \frac{1}{N}\sum_{i=1}^{N} x_i$ ($N \to \infty : \bar{x} \to \mu$ مقدار واقعی)

به عبارتی $\mu = \bar{x}_n \pm$ عدم قطعیت ($P\%$)

از طرفی وقتی تعداد مشاهدات در یک نمونه کم باشد (به خصوص $n<25$)، از طریق ضریب اصلاح بسل، از انحراف معیار تعدیل یافته (Adjusted Standard Deviation) به عنوان معیار عدم قطعیت استفاده می شود که به آن بهترین تخمین پراکندگی تکرار پذیری (Best Stimate of Precision) نیز گفته می شود.

$$s_n = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x}_n)^2}{n-1}} = \underbrace{\sqrt{\frac{n}{n-1}}}_{\text{Bessel correction factor}} \cdot \sigma_n$$

$n-1$ ← تعداد مشاهدات مستقل (درجه آزادی)

Bessel correction factor ($n\uparrow : \to 1$)

(با توجه به اینکه محاسبه میانگین هم توسط همین داده ها انجام شده، پس تعداد مشاهدات مستقل یکی کمتر است)

- وقتی تعداد داده های نمونه کم باشد (برخی منابع $n<60$ و برخی دیگر $n<200$ را مطرح نموده اند)، استفاده از توزیع Student's t برای تخمین سطح اطمینان و بازه عدم قطعیت مناسب تر و دقیق تر از محاسبه z می باشد. در این حالت، بازه تخمینی برای اندازه گیری کمیت مورد نظر عبارت است از:

$$x_i = \bar{x} \pm t_{\nu,p}\cdot s_x \quad (P\%)$$

(معادل $x_i = \mu \pm z\cdot\sigma \quad (P\%)$)

که در آن:

Student's t متغیر : $t = \frac{\bar{x}-\mu}{s_x/\sqrt{N}}$

Student's t distribution

ν	t_{50}	t_{90}	t_{95}	t_{99}
1	1.000	6.314	12.706	63.657
2	0.816	2.920	4.303	9.925
3	0.765	2.353	3.182	5.841
4	0.741	2.132	2.770	4.604
5	0.727	2.015	2.571	4.032
6	0.718	1.943	2.447	3.707
7	0.711	1.895	2.365	3.499
8	0.706	1.860	2.306	3.355
9	0.703	1.833	2.262	3.250
10	0.700	1.812	2.228	3.169
11	0.697	1.796	2.201	3.106
12	0.695	1.782	2.179	3.055
13	0.694	1.771	2.160	3.012
14	0.692	1.761	2.145	2.977
15	0.691	1.753	2.131	2.947
16	0.690	1.746	2.120	2.921
17	0.689	1.740	2.110	2.898
18	0.688	1.734	2.101	2.878
19	0.688	1.729	2.093	2.861
20	0.687	1.725	2.086	2.845
21	0.686	1.721	2.080	2.831
30	0.683	1.697	2.042	2.750
40	0.681	1.684	2.021	2.704
50	0.680	1.679	2.010	2.679
60	0.679	1.671	2.000	2.660
∞	0.674	1.645	1.960	2.576

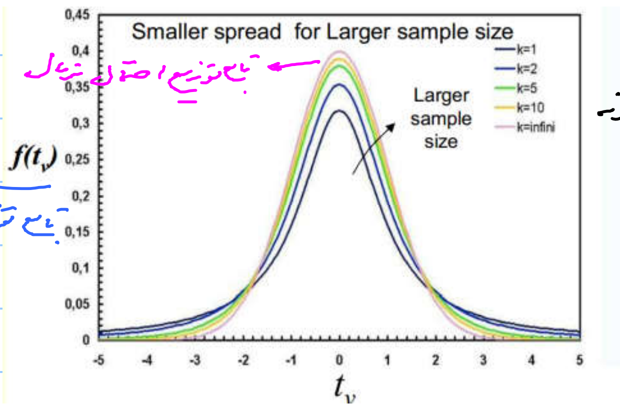
درجه آزادی — زیرنویس نیست

$$x_i = \underbrace{\bar{x}}_{\text{میانگین نمونه}} \pm \underbrace{t_{\nu,P}\, s_{\bar{x}}}_{\text{بازه عدم قطعیت}} \quad (P\%)$$

احتمال اینکه متغیر اندازه‌گیری شده در بازه مورد نظر قرار گیرد.

$$\int_{-t_\nu}^{t_\nu} f(t_\nu)\, dt_\nu = P\%$$

معادل z



تابع توزیع احتمال نرمال $\leftarrow f(t_\nu)$: $\infty \leftarrow k$ تعداد نمونه

- $\bar{x}_n$ (میانگین نمونه) چه میزان تخمین خوبی از میانگین واقعی (μ) است؟

با معرفی پارامتری به نام «انحراف معیار میانگین» می‌توان میزان عدم قطعیت میانگین نمونه از میانگین واقعی را تخمین زد.

انحراف معیار میانگین (standard deviation of the means) $s_{\bar{x}} = \frac{s_x}{\sqrt{N}}$

عدم قطعیت حول مقدار میانگین (با احتمال P)

$$\rightarrow \quad x' = \mu = x_{\text{true mean}} = \bar{x}_n \pm t_{\nu,P}\, s_{\bar{x}} \quad (P\%)$$

$$s_{\bar{x}} = \frac{s_x}{\sqrt{N}}$$

* حذف داده های پرت آماری (Data Outlier Detection / Rejection)

- داده های پرت آماری به دلایل مختلف می‌توانند در میان داده برداری های ما نمایان شوند. پس از به کارگیری معیارهای مناسب می‌توان آن ها را شناسایی و حذف نمود و مقادیر آماری را با داده های باقیمانده مجدداً محاسبه نمود.

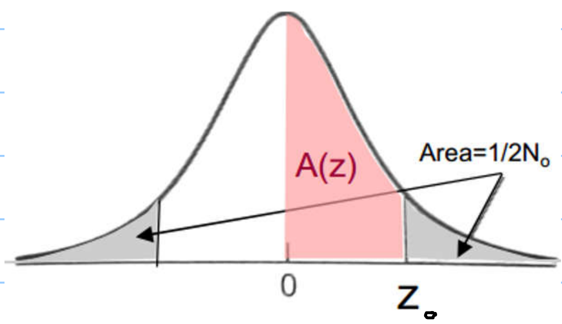
- یک راه سرانگشتی برای شناسایی داده های پرت آماری این است که اگر یک داده در نمونه‌ای بیش از 2σ از مقدار میانگین باشد، (با فرض توزیع نرمال) احتمال وقوعش کمتر از 5% می‌باشد و می‌تواند کنار گذاشته شود.

- معیار دیگر: معیار شوونه (Chauvenet's Criterion): اگر احتمال وقوع یک مقدار کمتر از $\frac{1}{2N_o}$ باشد (N_o تعداد داده ها/اندازه گیری ها)، آن داده قابل حذف است.

داده مشکوک(!)

$$Z_o = \left| \frac{x_i - \bar{x}}{\sigma_x \simeq S_x} \right| \xrightarrow{\text{جدول } z} \text{اگر } \left(1 - 2P(Z_o)\right) < \frac{1}{2N_o} \Rightarrow Z_o \text{ داده پرت است}$$

Number of Readings N_o	$1/(2N_o)$	z = Ratio of Max. acceptable deviation to Std. deviation= $\lvert x_i - \bar{x} \rvert / \sigma$
3	0.166	A(z)=(1-0.166)/2=0.417→z=1.38
4	0.125	1.54
5	0.1	1.65
10	0.05	1.96
25	0.02	2.33
50	0.01	2.57



مثال: در مجموعه داده زیر، کدام داده(ها) می‌توانند حذف شوند؟

x : 5.3, 5.73, 6.77, 5.26, 4.33, 5.45, 6.09, 5.64, 5.81, 5.75

$\bar{x}_{10} = 5.613$, $S_{10} = 0.627$

$N_o = 10 \rightarrow \frac{1}{2N_o} = 0.05 \rightarrow$ هر داده‌ای بیش از 1.96 انحراف معیار از میانگین فاصله داشته باشد با شوونه قابل حذف است

$\bar{x}_9 = 5.756$, $S_9 = 0.462$: در حالت جدید $\leftarrow$ پس طبق حکم $x_5 = 4.33$ را کنار می‌گذاریم

* رگرسیون یا برازش منحنی از روش حداقل مربعات (Curve Fitting Using Least Square Method)

- در دستگاه معادلات خطی $\underbrace{A_{n\times l}\,.\,\theta_{l\times 1}}_{}= b_{n\times 1}$ ، اگر A مربعی ($l=n$) و وارون پذیر

بردار مجهولات

باشد، داریم : $\theta = A^{-1}.b$

- اگر تعداد معادلات (n) بیش از تعداد مجهولات (l) باشد ($n>l$) ، برازش خطی

بهترین تخمین مجهولات ($\hat{\theta}$) از روش مینیمم مربعات (LSM) ، باید مجموع مربعات خطاها را کمینه نمود :

$$\hat{\theta}_{l\times 1} = \underset{\theta}{\operatorname{argmin}} \|E\|^2 = \underset{\theta}{\operatorname{argmin}} \underbrace{(A\theta - b)^T.(A\theta - b)}_{(E_1^2+E_2^2+\cdots+E_n^2)}$$

می توان نشان داد

$$\Rightarrow \hat{\theta}_{l\times 1} = \left((A^TA)^{-1}.A^T\right).b \quad (*)$$

بیان دیگر : $A\theta = b \Rightarrow \underbrace{A^TA}_{\text{ماتریس مربعی}}\theta = A^Tb \Rightarrow \theta = \underbrace{\left((A^TA)^{-1}A^T\right)}_{\text{pseudoinverse of A}}b$

- به صورت خاص فرض کنید که می خواهیم از مجموعه ای زوج داده ورودی - خروجی (x_i, y_i) ، بهترین چند جمله ای

درجه m عبور دهیم ؛ به عبارتی :
$$y_c = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m$$

که y_c تخمین مقدار واقعی خروجی به ازای ورودی x خواهد بود. هدف ، یافتن بردار مجهولات
$$\theta_{(m+1)\times 1} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix}$$

به ازای N زوج داده معلوم (x_i, y_i) ($N \geq m+1$) است :

مجموع مربع خطاهای تخمین از مقدار واقعی

$$D = \sum_{i=1}^{N}[y_i - (a_0 + a_1x + \cdots + a_mx^m)]^2 = \sum_{i=1}^{N} E_i^2$$

و با بهره گیری از فرم بسته (*)

$$dD = \frac{\partial D}{\partial a_0}da_0 + \frac{\partial D}{\partial a_1}da_1 + \cdots + \frac{\partial D}{\partial a_m}da_m$$

محاسبه مجهولات $a_0, a_1, \ldots, a_m$

$$\frac{\partial D}{\partial a_0} = 0 = \frac{\partial}{\partial a_0}\left\{\sum_{i=1}^{N}[y_i - (a_0 + a_1x + \cdots + a_mx^m)]^2\right\}$$

$$\frac{\partial D}{\partial a_1} = 0 = \frac{\partial}{\partial a_1}\left\{\sum_{i=1}^{N}[y_i - (a_0 + a_1x + \cdots + a_mx^m)]^2\right\}$$

$$\vdots$$

$$\frac{\partial D}{\partial a_m} = 0 = \frac{\partial}{\partial a_m}\left\{\sum_{i=1}^{N}[y_i - (a_0 + a_1x + \cdots + a_mx^m)]^2\right\}$$

$$\underbrace{\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ \vdots & & & & \\ 1 & x_N & x_N^2 & \cdots & x_N^m \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix}}_{\theta \text{ مجهولات}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}}_{b}$$

- در مواردی که تعداد زوج داده های موجود N بسیار زیاد باشد، از روش های عددی نظیر Gradient Descent برای یافتن بردار مجهولات $\underset{\sim}{\theta}$ (ضرایب بهترین منحنی عبوری) استفاده می‌شود.

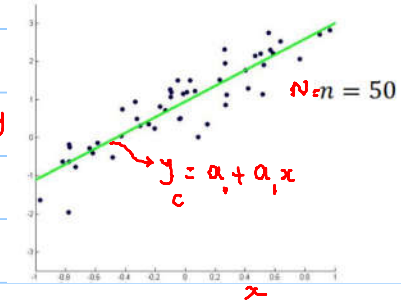
حدس مجهولات در گام t ام

نرخ آموزش

$$\underset{\sim}{\theta}^{t+1} = \underset{\sim}{\theta}^{t} - \eta \, \nabla_{\theta} D(\underset{\sim}{\theta}^{t})$$

- به صورت خاص، در حالت $m=1$ (بهترین خط عبوری از داده‌ها) داریم:

$$y_c = a_0 + a_1 x$$



$$\Rightarrow \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum_{K=1}^{N} x_K^2 & \sum x_K \\ \sum_{K=1}^{N} x_K & N \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum_{K=1}^{N} x_K y_K \\ \sum_{K=1}^{N} y_K \end{bmatrix}$$

$$\Rightarrow \begin{cases} a_1 = \dfrac{N \sum_{K=1}^{N} x_K y_K - \sum_{K=1}^{N} x_K \sum_{K=1}^{N} y_K}{N \sum_{K=1}^{N} x_K^2 - \left(\sum_{K=1}^{N} x_K\right)^2} = \dfrac{\text{کوواریانس ورودی-خروجی}}{\text{واریانس ورودی}} = \dfrac{S_{xy}^2}{S_{xx}^2} \quad (\ast\ast) \\[2ex] a_0 = \dfrac{\sum_{K=1}^{N} y_K \sum_{K=1}^{N} x_K^2 - \sum_{K=1}^{N} x_K \sum_{K=1}^{N} x_K y_K}{n \sum_{K=1}^{N} x_K^2 - \left(\sum_{K=1}^{N} x_K\right)^2} \end{cases}$$

$$(\ast\ast) \begin{cases} S_{xy}^2 = \dfrac{1}{N} \sum_{K=1}^{N} (y_K - \bar{y})(x_K - \bar{x}) \\[2ex] S_{xx}^2 = \dfrac{1}{N} \sum_{K=1}^{N} (x_K - \bar{x})^2 \end{cases}$$

$$a_1 \checkmark \Rightarrow \underline{a_0 = \bar{y} - a_1 \bar{x}} \checkmark$$

- برای برازش منحنی های غیرخطی، اگر بتوانیم رابطه غیرخطی را به شکل خطی تبدیل کنیم، استفاده از فرمول های ارائه شده همچنان امکان پذیر است:

الف) عبور منحنی کالیبراسیون توانی از داده ها (a, b مجهول) $\qquad y = a x^b$

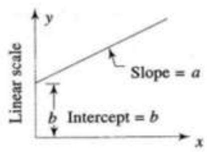
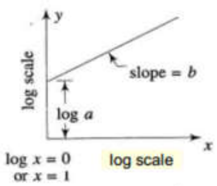
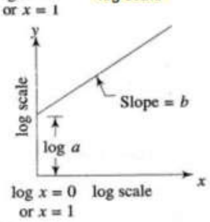
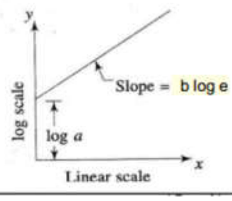
$$\rightarrow \log y_k = \log a + b \log x_k \qquad k = 1, 2, \dots, N$$

$$\Rightarrow \begin{bmatrix} 1 & \log x_1 \\ \vdots & \vdots \\ 1 & \log x_N \end{bmatrix} \begin{bmatrix} \log a \\ b \end{bmatrix} = \begin{bmatrix} \log y_1 \\ \vdots \\ \log y_N \end{bmatrix}$$

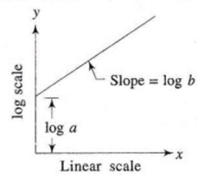
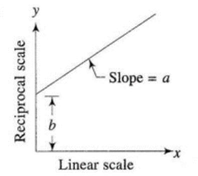
ب) عبور منحنی کالیبراسیون نمایی از داده ها: $\qquad y = a . b^x$

$$\rightarrow \log y_k = \log a + x_k . \log b \qquad k = 1, 2, \dots, N$$

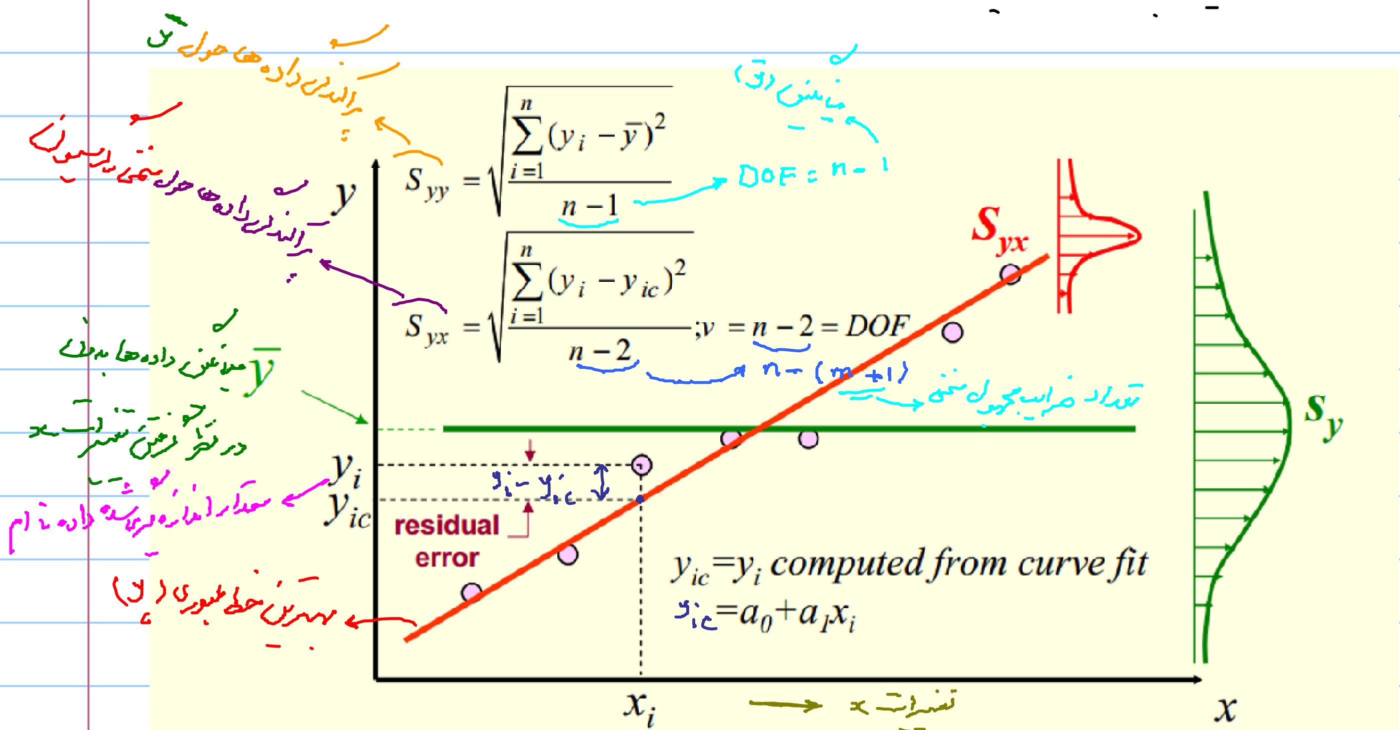
$$\Rightarrow \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix} \begin{bmatrix} \log a \\ \log b \end{bmatrix} = \begin{bmatrix} \log y_1 \\ \vdots \\ \log y_N \end{bmatrix}$$

Sl. No.	Type of curve and its governing equation	Ordinate	Abscissa	Recommended graph paper	Graphical representation showing slope and intercept parameter
1.	Linear or straight line $y = ax + b$	y	x	Liner versus linear	
2.	Power law curve $y = ax^b$ $y{=}a.x^b \Rightarrow \log y = \log a + b \log x$	$\log y$	$\log x$	log versus log	
3.	Modified power law curve $y = ax^b + c$ $y{-}c{=}a.x^b \Rightarrow \log(y{-}c) = \log a + b \log x$	$\log (y - c)$	$\log x$	log versus log	
4.	Exponential type curve $y = ae^{bx}$ $y{=}a.e^{bx} \Rightarrow \log y = \log a + b\,x\ \log e$	$\log y$	x	log versus linear	

(Contd.)

Sl. No.	Type of curve and its governing equation	Ordinate	Abscissa	Recommended graph paper	Graphical representation showing slope and intercept parameter
5.	Exponential type curve $y = ab^x$	$\log y$	x	log versus linear	
6.	Hyperbolic type curve $y = \frac{1}{ax + b}$	$\frac{1}{y}$	x	Reciprocal versus linear	

- انحراف معیار برای خط رگرسیون :



S_y : Spread around the mean ; $S_{y,x}$: Spread around the regression curve

Correlation coefficient: $r = \sqrt{1 - \dfrac{S_{yx}^2}{S_y^2}}$

ضریب همبستگی

- S_{yx} ، معرف پراکندگی / انحراف معیار داده‌ها حول منحنی رگرسیونی است که به آن بازه دقت برازش (precision interval) ، خطای تصادفی استاندارد برازش یا بازه اطمینان برازش (confidence interval) هم گفته می‌شود.

- احتمال اینکه نرخ داده (x_k, y_k) در بازه $y_c \pm t_{\nu,P}\frac{S_{yx}}{\sqrt{N}}$ قرار گیرد ، P% می‌باشد. (یعنی : احتمال P% ، $|y - y_{ck}|$ کوچکتر از $t_{\nu,P}\frac{S_{yx}}{\sqrt{N}}$ می‌باشد)

$y_c \pm \underbrace{t_{\nu,P}\frac{S_{yx}}{\sqrt{N}}}$ (P%) : منحنی رگرسیون و بازه اطمینان برازش با احتمال P%

عدم قطعیت به علت خطای تصادفی در متغیر وابسته (اندازه‌گیری شده) با احتمال P%

- ضریب همبستگی r ، معیاری از میزان شباهت و تطابق بین x و y است . حداکثر مقدار آن برابر 1 است و نشان دهنده انطباق کامل بین x و y است.

مثال: با توجه به داده‌های ارائه شده، بهترین رابطه خطی را بین x و y بیابید همچنین خطای استاندارد برازش (S_{yx}) و بازه اطمینان ٪۹۵ را محاسبه کنید.

x (cm)	y (V)
1.0	1.2
2.0	1.9
3.0	3.2
4.0	4.1
5.0	5.3

$$y_c = a_0 + a_1 x$$

از روابط ارائه شده برای بهترین خط عبوری $\longrightarrow a_1 = 1.04 \quad , \quad a_0 = 0.02 \Rightarrow y_c = 1.04x + 0.02$

$$S_{yx} = \sqrt{\frac{\sum_{i=1}^{5} (y_i - y_{ci})}{\underset{\nu \swarrow \quad \nwarrow m+1}{N-2}}} \Rightarrow \underline{S_{yx} = 0.16}$$

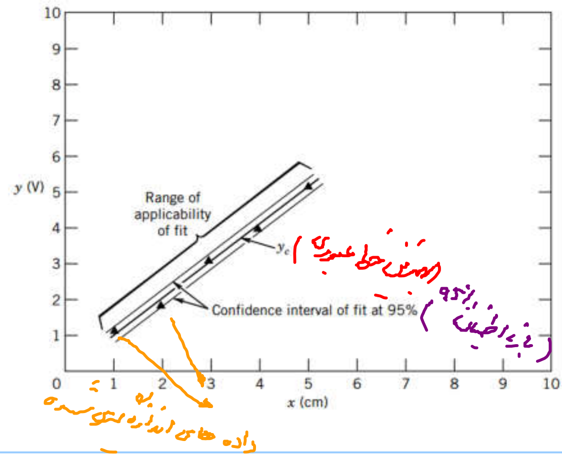
$$\Rightarrow t_{3,95} \cdot \frac{S_{yx}}{\sqrt{5}} = 0.23$$

عدم قطعیت تصادفی
با احتمال ٪۹۵ منحنی برازش

$$t_{3,95} = 3.18$$

$3 \to \nu$, $95 \to P\%$

$$\Rightarrow y_c = 1.04x + 0.02 \pm 0.23 \ \text{V} \ (95\%)$$



مثال: یک سنسور اندازه‌گیری، ولتاژ متناسب با سرعت را به صورت $E = a + bU^m$ خروجی می‌دهد. با توجه به داده‌های ارائه شده، بهترین منحنی عبوری و بازه اطمینان ٪۹۵ را بیابید.

U (m/s)	E (V)
0.0	3.19
10.0	3.99
20.0	4.30
30.0	4.48
40.0	4.65

مجهول: a, b, m

در سرعت $U = 0$ ، ولتاژ اندازه‌گیری شده 3.197 است. بنابراین $a = 3.19$.

$$E = a + bU^m \xrightarrow{\text{تبدیل به فرم خطی}} \underbrace{\log(E-a)}_{Y} = \underbrace{\log b}_{B} + m\,\underbrace{\log U}_{X}$$

داده $N = 4$

U	$Y = log(E\text{-}a)$	$X = log\ U$	$E_i - E_{c_i}$
0.0	—	—	0.0
10.0	−0.097	1.000	−0.01
20.0	0.045	1.301	0.02
30.0	0.111	1.477	0.0
40.0	0.164	1.602	−0.01

از روابط بهترین خط مستقیم $\Rightarrow$

$$\begin{cases} B = \log b = -0.525 \\ m = 0.43 \end{cases} \Rightarrow Y = 0.43X - 0.525$$

$$S_{yx} = \sqrt{\frac{\sum_{i=1}^{4}(Y_i - Y_{c_i})}{\nu}} = 0.007 \Rightarrow$$

بازه عدم قطعیت با اطمینان 95٪ : $t_{2,95} \cdot \frac{S_{yx}}{\sqrt{N}} = 0.015$

$t_{2,95} = 4.30$, $\nu = 4 - 2 = 2$, $N = 4$

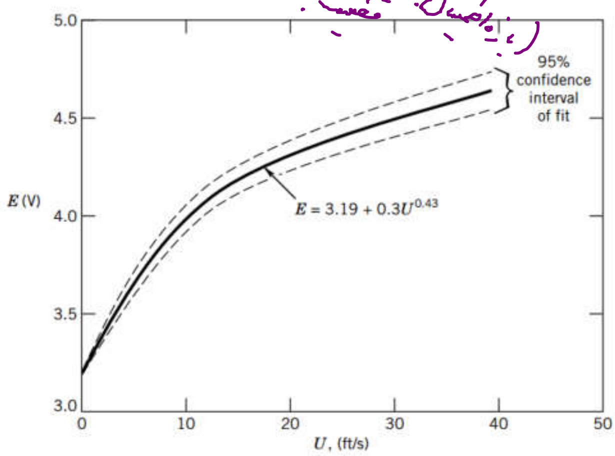
$$\Rightarrow Y = 0.43X - 0.525 \pm 0.015 \quad (P = 95\%)$$

(ناحیه اطمینان 95٪ برای Fit منحنی)

$$\log(E-a) = \log b + m \log U$$

$$E = a + b \cdot U^m$$

$$\Rightarrow E = 3.19 + 0.30\,U^{0.43}$$



* انتشار خطا (Uncertainty Propagation) در کمیت‌های دیگر

- می‌خواهیم خطا در کمیت وابسته (dependent) y را دارای خطا در متغیرهای مستقل (independent) x_i محاسبه نماییم: (مثال: وابستگی اختلاف سطح ستون جیوه به پارامترهای مختلف)

$$y = f(x_1, x_2, \dots, x_n)$$

عدم قطعیت در x_1

$$(\text{خطا در } y)\ \Delta y = f(x_1+\delta x_1, x_2+\delta x_2, \dots, x_n+\delta x_n) - f(x_1, x_2, \dots, x_n)$$

بسط سری تیلور — تا ترتیب اول

$$\Delta y = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\,\delta x_i \Longrightarrow \boxed{\text{حداکثر خطای ممکن: } \Delta y \leqslant \sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i}\,\delta x_i \right|}$$

- می‌توان نشان داد که انحراف معیار y برحسب عدم قطعیت‌ها / انحراف معیارهای x_i به شکل زیر است،

$$\sigma_y = \left(\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)^2 \sigma_x^2 \right)^{1/2}$$

- ممکن است در طراحی‌ها بخواهیم که حد خطا روی x_i ها را به گونه‌ای تعیین کنیم که خطا روی y از میزان معینی کمتر شود:

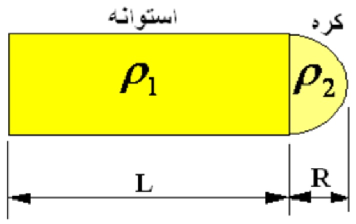
$$|\Delta y|_{max} = \sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i} \right| |\delta x_i| \leqslant \delta_{y_D} \text{ : مقدار مطلوب خطا} \Rightarrow \boxed{|\delta x_i| \leqslant \frac{\delta y_D}{n \left| \frac{\partial f}{\partial x_i} \right|}}$$

به طور مشابه

$$|\sigma_y|^2_{y,max} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)^2 \sigma_{x_i}^2 \leqslant \sigma_{y_D}^2 \rightarrow \sigma_{x_i} \leqslant \frac{\sigma_{y_D}}{\sqrt{n} \left| \frac{\partial f}{\partial x_i} \right|}$$

- در روابط جبری فوق عدم قطعیت‌ها، δx_i ها همگی با یک سطح اطمینان (P%) استفاده شوند.

سوال: در شکل زیر، جرم جسم و میزان عدم قطعیت آن را محاسبه نمایید.



$L = 100 \pm 0.1\ cm \quad \rho_1 = 3.5 \pm 0.1\ g/cm^3$

$R = 4 \pm 0.05\ cm \quad \rho_2 = 2.5 \pm 0.05\ g/cm^3$

$\bar{M}_{کل} = ? \quad , \quad \delta M = ?$

$$M_{کل} = \rho_1 (\pi R^2) L + \rho_2 \cdot \left(\frac{2}{3}\pi R^3\right) = f(\rho_1, R, L, \rho_2)$$

عدم قطعیت در جرم

$$\delta M = (\pi \bar{R}^2)\bar{L}\,\delta\rho_1 + \bar{\rho}_1(\pi\bar{R}^2)\,\delta L + 2\bar{\rho}_1(\pi\bar{R}\,\delta R)\bar{L} + \frac{2}{3}\pi\bar{R}^3\,\delta\rho_2 + 2\bar{\rho}_2\pi\bar{R}^2\,\delta R$$

$$\Rightarrow \sigma_M = \left(\left((\pi\bar{R}^2)\bar{L}\right)^2\sigma_{\rho_1}^2 + \left(\bar{\rho}_1(\pi\bar{R}^2)\right)^2\sigma_L^2 + \left(\frac{2}{3}\pi\bar{R}^3\right)^2\sigma_{\rho_2}^2 + \left(2\bar{\rho}_2\pi\bar{R}^2 + 2\bar{\rho}_1(\pi\bar{R})\bar{L}\right)^2\sigma_R^2 \right)^{1/2}$$

جایگذاری مقادیر در روابط $\rightarrow$ $\bar{M} = 2094.4\ g \quad , \quad \delta_M = 77.97\ g$

سوال: قطر نمونه‌ای شافت و بوش یاتاقان آن به صورت زیر داده شده است:

$\bar{D}_s = 31\ mm \quad , \quad \sigma_s = 0.04\ mm$ $\rightarrow$ Shaft diameter & SD

$\bar{D}_h = 31.1\ mm \quad , \quad \sigma_h = 0.03\ mm$ $\rightarrow$ hub diameter & SD

در صورتی که لقی $(C = D_h - D_s)$ مجاز در بازه 0.18 تا 0.03 میلیمتر باشد، چند درصد از شافت‌ها و بوش‌ها مناسب نبوده و باید برگشت بخورند؟

Clearance $= C = D_h - D_s = f(D_h, D_s)$

$\bar{C} = \bar{D}_h - \bar{D}_s \Rightarrow \bar{C} = 0.1\ mm$

$$\sigma_c = \left(\left(\frac{\partial c}{\partial D_h}\cdot\sigma_{D_h}\right)^2 + \left(\frac{\partial c}{\partial D_s}\cdot\sigma_{D_s}\right)^2 \right)^{1/2} \Rightarrow \sigma_c = 0.05\ mm$$

- احتمال اینکه یک حبه شکلات و پوشش در محدوده مناسب باشد:

فرض توزیع نرمال

$$P(0.03 < C < 0.18) = P\left(\frac{0.03-0.1}{0.05} < \underbrace{\frac{C-\bar{C}}{\sigma_C}}_{Z} < \frac{0.18-0.1}{0.05}\right)$$

Z محاسبه

$$= P(-1.4 < Z < 1.6) = \underbrace{\Phi(1.6)}_{0.9452} - \underbrace{\Phi(-1.4)}_{1-\Phi(1.4)} = 0.8644$$

$$1-\underbrace{\Phi(1.4)}_{0.9192}$$

- بنابراین احتمال اینکه یک حبه شکلات و پوشش مناسب نباشد، عبارت است از

$$1 - P(0.03 < C < 0.18) = 1 - 0.8644 = \underline{13.6\,\%}$$