Scientific Measurement

PHY-4103

Dr. Eram Rizvi & Dr. Alston Misquitta

Lecture 6 - The Gaussian Probability Distribution

Queen Mary University of London

Last lecture we looked at the Binomial probability distribution:

$$f(k;n,p) = \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k}$$

$$= {}_nC_k p^k (1-p)^{n-k}$$

$$= {}_nC_k p^k q^{n-k} \quad \text{where } q = 1-p \quad \text{q=P(fail)}$$

Prob. of k successes given n trials,
and probability of success for single trial = p

Used to determine uncertainties in situations where each trial has
a true/false or yes/no result, e.g. opinion polls & efficiency calcs.

Another common probability distribution: The Poisson Distribution

Relevant for counting experiments

Describes statistics of events occurring at a random but at a definite rate e.g.

number of radioactive decays in 1 min interval

number of births in UK per 24 hours

Experiment: count number of radioactive decays in 1 min interval: k

Assume 100% efficient detector - no error on k

Repeat experiment → get different k ⇒ statistical fluctuation!

n nuclei	$n \sim 10^{20}$
p prob of single decay in 1 min	$p \sim 10^{-20}$
k decays measured in 1 min	$n \times p \sim$ finite

Can use binomial distribution in simplified form ⇒ Poisson Distribution

As $n \to \infty$ and $p \to 0$ with $np \to \lambda$ (const) then binomial dist gives Poisson dist.

Probability of observing k counts if mean number of counts is λ

$$f(\lambda,k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$\lambda = n \times p$ = mean no. of counts

k = observed no. of counts

Important properties of the Poisson distribution:

$$\lambda = \bar{k}$$

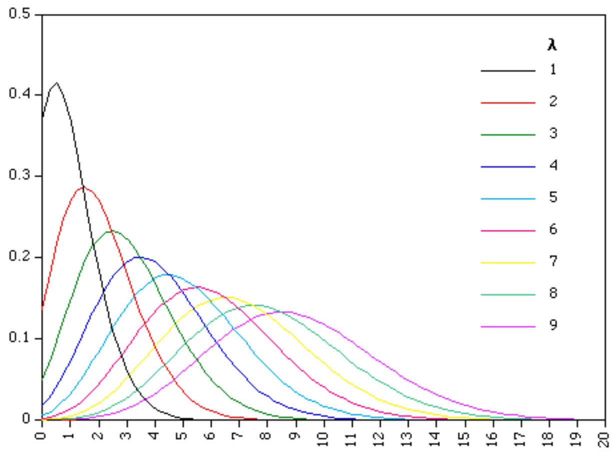
$$\sigma = \sqrt{\lambda}$$

In other words:

measurement of k gives estimate of λ

measurement of $\sqrt{k}$ gives estimate of σ

i.e. the <u>true</u> mean is given by the mean of k and also k gives the true variance σ^2



Poisson dist. skewed for small λ

More symmetric as λ increases

1 measurement gives λ and σ

$$k \pm \sqrt{k}$$

fractional error: $\frac{\sigma}{\lambda} = \frac{\sqrt{k}}{k} = \frac{1}{\sqrt{k}}$

Thus error decreases as k increases but only as the square root!

The most important probability distribution: Gaussian

Other names: Bell curve, Normal Distribution

Binomial & Poisson dists. are discrete
Gaussian is continuous probability distribution

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

μ = mean of distribution
σ = standard deviation
x = measured variable

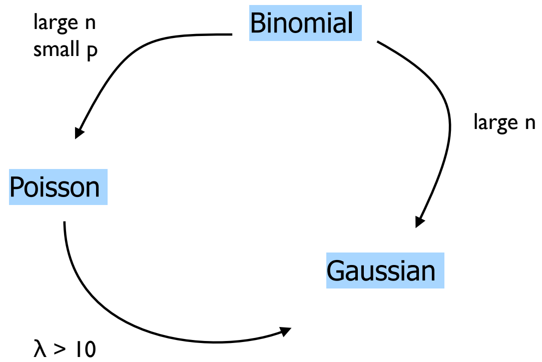
normalisation factor $\int_{-\infty}^{+\infty} G(x)dx = 1$

Gaussian distribution has nice properties:

mean = median = mode: $\mu = \int_{-\infty}^{+\infty} xG(x)dx = 1$

$$\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 G(x)dx$$

- Gaussian dist. is symmetric about mean μ
- μ shifts the peak of distribution



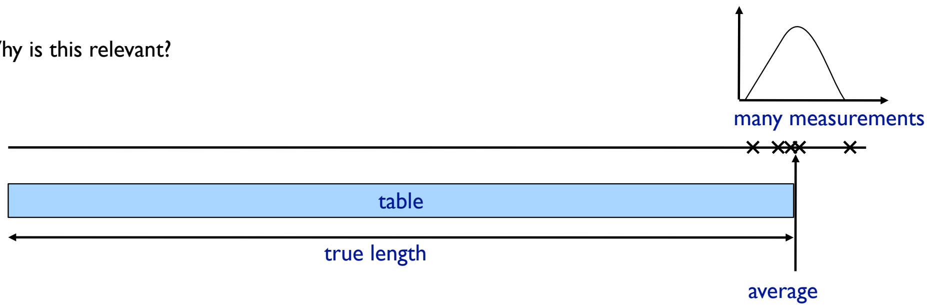
Why is the Gaussian Distribution so important?

The Central Limit Theorem

The mean of a large number of independent samples each of size n from the same probability distribution (not necessarily Gaussian) has approximately a Gaussian distribution centred on the population mean and variance which reduces as 1/n

http://www.intuitor.com/statistics/CLAppClasses/CentLimApplet.htm

Why is this relevant?



Return to the measurement of our table
Repeat this 100 times
We will see a spread of measurements
Spread arises from many small random effects
Central Limit Theorem tells us that the spread will be Gaussian

This explains why Gaussian errors appear everywhere!

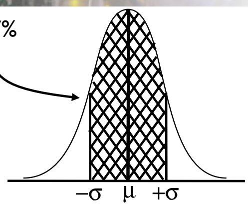
Confidence Intervals

38 Theoretical distributions Chap. 3

TABLE 3.2
TWO-TAILED GAUSSIAN INTEGRAL
Giving the percentage probability that a point lies within the given number of σ from the mean

$\frac{x-\mu}{\sigma}$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.00	0.80	1.60	2.39	3.19	3.99	4.78	5.58	6.38	7.17
0.10	7.97	8.76	9.55	10.34	11.13	11.92	12.71	13.50	14.28	15.07
0.20	15.85	16.63	17.41	18.19	18.97	19.74	20.51	21.28	22.05	22.82
0.30	23.58	24.34	25.10	25.86	26.61	27.37	28.12	28.86	29.61	30.35
0.40	31.08	31.82	32.55	33.28	34.01	34.73	35.45	36.16	36.88	37.59
0.50	38.29	38.99	39.69	40.39	41.08	41.77	42.45	43.13	43.81	44.48
0.60	45.15	45.81	46.47	47.13	47.78	48.43	49.07	49.71	50.35	50.98
0.70	51.61	52.23	52.85	53.46	54.07	54.67	55.27	55.87	56.46	57.05
0.80	57.63	58.21	58.78	59.35	59.91	60.47	61.02	61.57	62.11	62.65
0.90	63.19	63.72	64.24	64.76	65.28	65.79	66.29	66.80	67.29	67.78
1.00	68.27	68.75	69.23	69.70	70.17	70.63	71.09	71.54	71.99	72.43
1.10	72.87	73.30	73.73	74.15	74.57	74.99	75.40	75.80	76.20	76.60
1.20	76.99	77.37	77.75	78.13	78.50	78.87	79.23	79.59	79.95	80.29
1.30	80.64	80.98	81.32	81.65	81.98	82.30	82.62	82.93	83.24	83.55
1.40	83.85	84.15	84.44	84.73	85.01	85.29	85.57	85.84	86.11	86.38
1.50	86.64	86.90	87.15	87.40	87.64	87.89	88.12	88.36	88.59	88.82
1.60	89.04	89.26	89.48	89.69	89.90	90.11	90.31	90.51	90.70	90.90
1.70	91.09	91.27	91.46	91.64	91.81	91.99	92.16	92.33	92.49	92.65
1.80	92.81	92.97	93.12	93.28	93.42	93.57	97.71	93.85	93.99	94.12
1.90	94.26	94.39	94.51	94.64	94.76	94.88	95.00	95.12	95.23	95.34
2.00	95.45	95.56	95.66	95.76	95.86	95.96	96.06	96.15	96.25	96.34
2.10	96.43	96.51	96.60	96.68	96.76	96.84	96.92	97.00	97.07	97.15
2.20	97.22	97.29	97.36	97.43	97.49	97.56	97.62	97.68	97.74	97.80
2.30	97.86	97.91	97.97	98.02	98.07	98.12	98.17	98.22	98.27	98.32
2.40	98.36	98.40	98.45	98.49	98.53	98.57	98.61	98.65	98.69	98.72

Area=68.27%



Total area under the curve = 1
Normalisation condition!
Integrating part of the curve tells us of probability
E.g. prob of result within ±σ

68.3% within ±1σ
95.5% within ±2σ
99.7% within ±3σ

Thus measurement quoted as

63 ± 1 cm means

68% prob within 62-64 cm
95% prob within 61-65 cm
99% prob within 60-66 cm

TABLE 3.3
ONE-TAILED GAUSSIAN INTEGRAL
Giving the percentage probability that a point lies within the given number of σ to one side of the mean

$\frac{x-\mu}{\sigma}$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	50.00	50.40	50.80	51.20	51.60	51.99	52.39	52.79	53.19	53.59
0.10	53.98	54.38	54.78	55.17	55.57	55.96	56.36	56.75	57.14	57.53
0.20	57.93	58.32	58.71	59.10	59.48	59.87	60.26	60.64	61.03	61.41
0.30	61.79	62.17	62.55	62.93	63.31	63.68	64.06	64.43	64.80	65.17
0.40	65.54	65.91	66.28	66.64	67.00	67.36	67.72	68.08	68.44	68.79
0.50	69.15	69.50	69.85	70.19	70.54	70.88	71.23	71.57	71.90	72.24
0.60	72.57	72.91	73.24	73.57	73.89	74.22	74.54	74.86	75.17	75.49
0.70	75.80	76.11	76.42	76.73	77.04	77.34	77.64	77.94	78.23	78.52
0.80	78.81	79.10	79.39	79.67	79.95	80.23	80.51	80.78	81.06	81.33
0.90	81.59	81.86	82.12	82.38	82.64	82.89	83.15	83.40	83.65	83.89
1.00	84.13	84.38	84.61	84.85	85.08	85.31	85.54	85.77	85.99	86.21
1.10	86.43	86.65	86.86	87.08	87.29	87.49	87.70	87.90	88.10	88.30
1.20	88.49	88.69	88.88	89.07	89.25	89.44	89.62	89.80	89.97	90.15
1.30	90.32	90.49	90.66	90.82	90.99	91.15	91.31	91.47	91.62	91.77
1.40	91.92	92.07	92.22	92.36	92.51	92.65	92.79	92.92	93.06	93.19
1.50	93.32	93.45	93.57	93.70	93.82	93.94	94.06	94.18	94.29	94.41
1.60	94.52	94.63	94.74	94.84	94.95	95.05	95.15	95.25	95.35	95.45
1.70	95.54	95.64	95.73	95.82	95.91	95.99	96.08	96.16	96.25	96.33
1.80	96.41	96.49	96.56	96.64	96.71	96.78	96.86	96.93	96.99	97.06
1.90	97.13	97.19	97.26	97.32	97.38	97.44	97.50	97.56	97.61	97.67
2.00	97.72	97.78	97.83	97.88	97.93	97.98	98.03	98.08	98.12	98.17
2.10	98.21	98.26	98.30	98.34	98.38	98.42	98.46	98.50	98.54	98.57
2.20	98.61	98.64	98.68	98.71	98.75	98.78	98.81	98.84	98.87	98.90
2.30	98.93	98.96	98.98	99.01	99.04	99.06	99.09	99.11	99.13	99.16
2.40	99.18	99.20	99.22	99.25	99.27	99.29	99.31	99.32	99.34	99.36

Can integrate up to some max value
One tailed Gaussian
integral = 0.5 + ½ (area $-\sigma$ to $+\sigma$)
84.1% < 1σ
97.7% < 2σ