FOUNDATION YEAR 2025-2026

COURSE - DATA ANALYSIS

Lesson 6

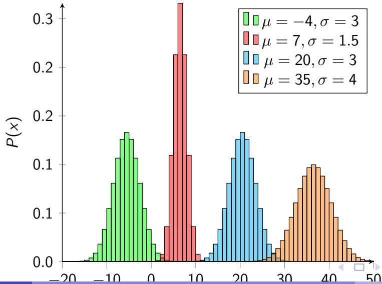
Niccolò Maffezzoli, Ca' Foscari University of Venice

January 29, 2026

Gaussian Distribution

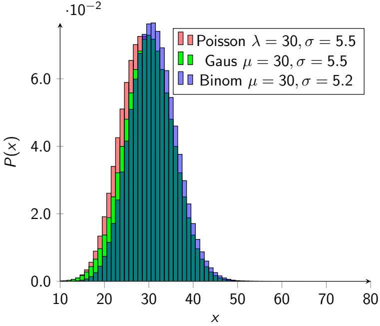
$$P(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

- The Binomial and Poisson distribution are discrete probability distributions. The Gaussian distribution is a continuous probability distribution. Negative values are fine.
- The distribution is symmetrical wrt to μ (mean/median/mode).
- The width is controlled by σ.



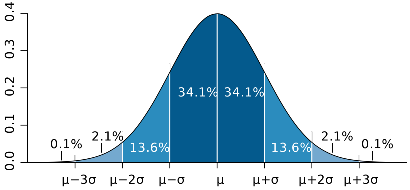
Gaussian Distribution

- A Poisson with a large λ tends to a Gaussian shape with $\mu = \lambda$ and $\sigma = \sqrt{\lambda}$. 'Large' is roughly $\lambda \approx 5 - 10$, depending on the desired accuracy.
- The Binomial with large μ tends to a Gaussian with $\mu = np$ and $\sigma = \sqrt{np(1-p)}$. 'Large' is roughly $np \approx 5$.



Gaussian Distribution - Integrals

You measure the length of a table and you find the length (X) measurements to be distributed according to a Gaussian (with $\sigma = 1$). What is the probability that any measurement is exactly equal to the mean ?



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

$$= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\mu-\mu)^2}{2\sigma^2}} = \frac{1}{\sigma\sqrt{2\pi}} \simeq 0.4$$

In practice the question should rather be: what's the probability that my measurement is close to the mean by a certain amount, e.g. 0.7 ? P($\mu - 0.7 < X < \mu + 0.7$)? Or what is the probability that a measurement lies between x1 and x2? P($x_1 < X < x_2$) ? Or P($X > x_3$) ? Or P($X < x_4$) ?

Gaussian Distribution - Integrals

Gaussian integrals are tabulated (or can be calculated by computers).

Procedure to evaluate Gaussian Integrals

1. Convert x to $z = (x - \mu)/\sigma$ (distance of x from the mean, in terms of σ).
2. Integrals are tabulated with respect to z.
3. Table 3.2: probability that a point lies between -z and +z.
4. Remember that the integral of the Gaussian is 1 (like all probability distributions).

Gaussian Distribution - Integral Tables (page 38 Barlow)

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
2.50	98.76	98.79	98.83	98.86	98.89	98.92	98.95	98.98	99.01	99.04
2.60	99.07	99.09	99.12	99.15	99.17	99.20	99.22	99.24	99.26	99.29
2.70	99.31	99.33	99.35	99.37	99.39	99.40	99.42	99.44	99.46	99.47
2.80	99.49	99.50	99.52	99.53	99.55	99.56	99.58	99.59	99.60	99.61
2.90	99.63	99.64	99.65	99.66	99.67	99.68	99.69	99.70	99.71	99.72
3.00	99.73	99.74	99.75	99.76	99.76	99.77	99.78	99.79	99.79	99.80
3.10	99.81	99.81	99.82	99.83	99.83	99.84	99.84	99.85	99.85	99.86
3.20	99.86	99.87	99.87	99.88	99.88	99.88	99.89	99.89	99.90	99.90
3.30	99.90	99.91	99.91	99.91	99.92	99.92	99.92	99.92	99.93	99.93
3.40	99.93	99.94	99.94	99.94	99.94	99.94	99.94	99.95	99.95	99.95
3.50	99.95	99.96	99.96	99.96	99.96	99.96	99.96	99.96	99.97	99.97

Gaussian Distribution - Exercise

Exercise

X is a variable distributed according to a Gaussian Distribution with $\mu = 8$, $\sigma = 2$. What is the probability that a value x is greater than 10.46? $P(x > 10.46)$?

1. Transform x to z. $z = (x - \mu)/\sigma = (10.46 - 8)/2 = 1.23$
2. From the integral tables we read that $P(-1.23 < z < 1.23) = 0.7813$
3. $P(z > 1.23) = (1 - 0.7813)/2 = 0.10935$

What is the probability that a value x is greater than 12.86? $P(x > 12.86)$?

1. Transform x to z. $z = (x - \mu)/\sigma = (12.86 - 8)/2 = 2.43$
2. From the integral tables we read that $P(-2.43 < z < 2.43) = 0.9849$
3. $P(z > 2.43) = (1 - 0.9849)/2 = 0.00755$

Exercise 3.3 page 47

During a meteor shower, meteors fall at the rate 15.7 per hour.

- What is the probability of observing less than 5 in a given period of 30 minutes?
- Bonus: solve the exercise in python using the function you defined last time.
- Bonus-bonus: .. and using a for loop.

Exercise 3.3 page 47

During a meteor shower, meteors fall at the rate 15.7 per hour.

- What is the probability of observing less than 5 in a given period of 30 minutes?
- Bonus: solve the exercise in python using the function you defined last time.
- Bonus-bonus: .. and using a for loop.

Solution

Poisson distribution with $\lambda = 7.85\ meteors/30m$

$P(X \leq 4) = \sum_{k=0}^{4} \frac{\lambda^k e^{-\lambda}}{k!} \approx 10.9\%$

Exercise 2

Construct a table giving the probability distribution for the number of 3's seen in 7 rolls of a standard die. Find the mean and standard deviation of this probability distribution. Check the values using the python function you defined.

Exercise 2

Construct a table giving the probability distribution for the number of 3's seen in 7 rolls of a standard die. Find the mean and standard deviation of this probability distribution. Check the values using the python function you defined.

Solution

Binomial distribution: $P(k, p, n) = p^k (1-p)^{n-k} \frac{n!}{k!(n-k)!}$
In our case, $p = 1/6$, $n = 7$. You need to calculate P(k=0), P(k=1), P(k=2), ... P(k=7).
$\mu = n \cdot p = 7/6 \approx 1.17$. $\sigma^2 = n \cdot p \cdot q = 7 \cdot 1/6 \cdot 5/6 = 35/36$. $\sigma = \sqrt{35/36}$.