Distribusi Peluang Kontinu

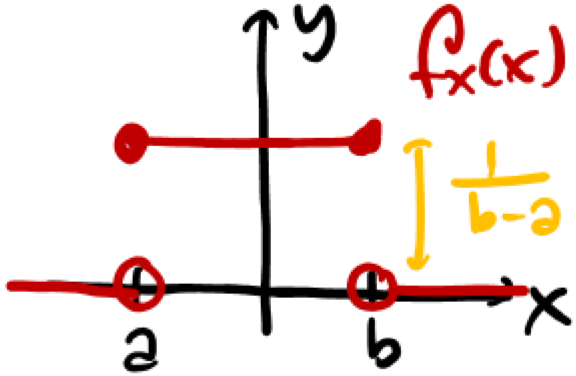
Distribusi Uniform, Normal, dan Eksponensial

2.1 Distribusi Uniform, Normal, dan Eksponensial

	Uniform	Normal	Eksponensial
Karakteristik distribusi	Setiap nilai dalam interval memiliki kemungkinan yang sama untuk muncul.	Distribusi peluang yang ideal dan simetris, dengan rata-rata, median, dan modus yang bernilai sama.	Biasanya digunakan untuk menggambarkan waktu tunggu hingga kejadian Poisson terjadi dan memiliki sifat *memoryless*.
Notasi	$X \sim U(a, b), \quad a, b \in \mathbb{R}$	$X \sim \mathcal{N}(\mu, \sigma), \quad \mu \in \mathbb{R}, \sigma \in \mathbb{R}^+$	$X \sim \text{Exp}(\lambda), \quad \lambda \in \mathbb{R}^+$
Support	$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$ ✓	$(-\infty, \infty) := \mathbb{R}$	$[0, \infty) := \{x \in \mathbb{R} \mid x \geq 0\}$
PDF: $f_X(x)$	$\frac{1}{b-a}, \quad a \leq x \leq b$ ✓ (konstan)	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\lambda e^{-\lambda x}$
CDF: $F_X(x)$	$\frac{x-a}{b-a}$	$\int_{-\infty}^{x} f_X(t)\, dt$	$1 - e^{-\lambda x}$

batas bawah: a, batas atas: b.

CALVIN INSTITUTE OF TECHNOLOGY

	Uniform	Normal	Eksponensial
Sketsa PDF			
✓ Ekspektasi	$\frac{a+b}{2}$	μ	$\frac{1}{\lambda}$
✓ Variansi	$\frac{b^3 - 3b^2a + 3ba^2 - a^3}{12(b-a)}$	σ^2	$\frac{1}{\lambda}$

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & x \text{ lainnya} \end{cases} \rightsquigarrow \boxed{\mathbb{E}[X] = \frac{a+b}{2}}$$

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx = \int_a^b x \cdot \frac{1}{b-a}\, dx = \frac{1}{b-a} \int_a^b x\, dx$$

$$= \frac{1}{b-a} \left[\frac{x^2}{2}\right]_a^b = \frac{1}{b-a} \left[\frac{b^2}{2} - \frac{a^2}{2}\right] = \frac{1}{b-a} \cdot \frac{b^2-a^2}{2}$$

$$= \frac{1}{\cancel{b-a}} \frac{\cancel{(b-a)}(b+a)}{2} = \frac{a+b}{2} \quad //$$

Suatu ruang pertemuan disyaratkan dapat digunakan untuk rapat dengan durasi paling lama dua jam. Lamanya durasi suatu rapat diasumsikan berdistribusi uniform. Tuliskan PDF-nya untuk menentukan berapakah peluang sebuah rapat berlangsung tidak lebih dari 45 menit?

120 menit

Variabel acak

Misalkan X menyatakan durasi rapat. $\rightarrow S_X = [0, 120]$.

$$\underset{\substack{\downarrow \\ \text{batas bawah}}}{X \sim U(0}, \underset{\substack{\downarrow \\ \text{batas atas}}}{120)} \longrightarrow f_X(x) = \begin{cases} \dfrac{1}{120-0} = \dfrac{1}{120}, & x \in [0, 120] \\ 0, & x \text{ lainnya.} \end{cases}$$

- Ditanya: $\mathbb{P}(\text{rapat tdk lebih dari 45 menit}) = \mathbb{P}(X \leq 45)$

$$\mathbb{P}(X \leq 45) = F_X(45) \quad \text{(CDF)}$$

$$= \frac{45-0}{120-0} = \frac{45}{120} \;//$$

$$= \int_0^{45} \frac{1}{120}\, dx \quad f_X(x)$$

$$= \left[\frac{x}{120}\right]_0^{45} = \frac{45}{120} \;//$$

2.2 Eksplorasi Distribusi Normal

Apa yang membedakan distribusi uniform dengan distribusi normal?

- Dist. Uniform: ______________ *random*.
- Dist. Normal: ______________ *random*.

Definisi. Distribusi normal baku adalah distribusi normal dengan $\mu =$ _____ dan $\sigma =$ _____. Kita notasikan Z berdistribusi normal baku sebagai $Z \in \mathcal{N}(0, 1)$. Dengan demikian, kita peroleh fungsi kepadatan peluang dan fungsi distribusi kumulatifnya adalah

$$\varphi(z) := \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \qquad \Phi(x) := \mathbb{P}(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt.$$

Teorema. Misalkan $X \sim \mathcal{N}(\mu, \sigma)$ dengan transformasi $Z =$ ______, maka $Z \sim \mathcal{N}(0, 1)$ sehingga

Misalkan peubah acak $X \sim \mathcal{N}(15, 9)$, hitung nilai dari $\mathbb{P}(X < 20)$.

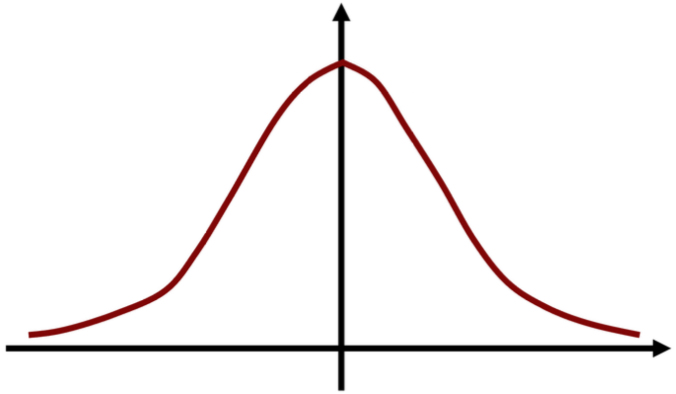
Misalkan peubah acak $X \sim \mathcal{N}(-17, 20)$, hitung nilai dari $\mathbb{P}(-16 < X < 1)$.

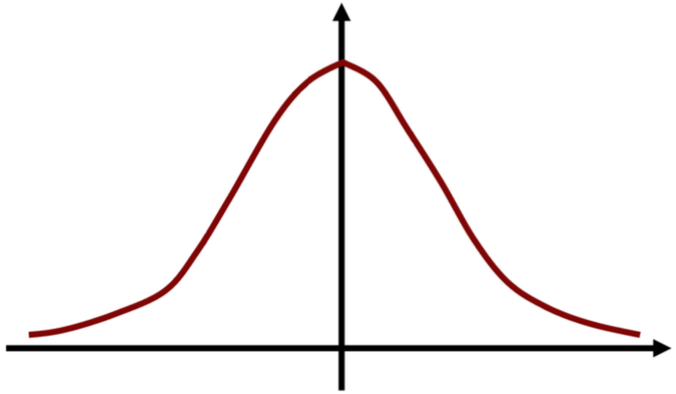
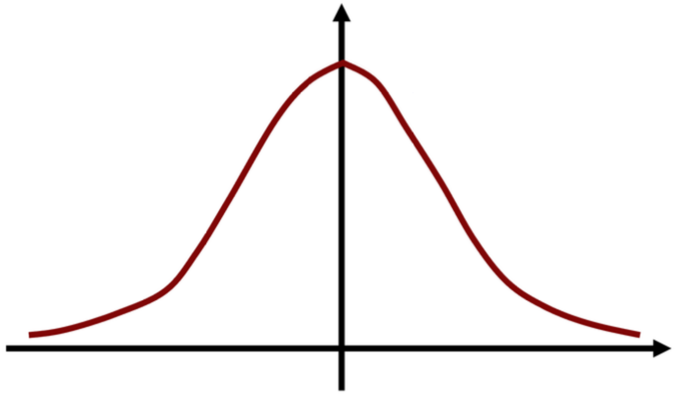
(a)

$$\mathbb{P}(X < x) = \mathbb{P}\left(Z < \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right),$$

(b)

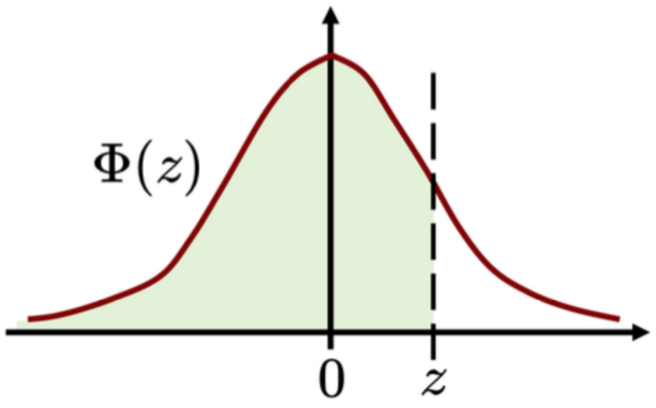
$$\mathbb{P}(a < X < b) = \mathbb{P}\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Jenis kejadian	Perhitungan dengan $\Phi(z)$	Visualisasi luas daerah di bawah kurva
$\mathbb{P}(Z < z)$		

Jenis kejadian	Perhitungan dengan $\Phi(z)$	Visualisasi luas daerah di bawah kurva
$\mathbb{P}(Z > z)$		
$\mathbb{P}(z_1 < Z < z_2)$		

(a) Tinggi badan mahasiswa CIT didata dan ditemukan bahwa datanya berdistribusi normal dengan rata-rata 165 cm dan standar deviasi 18 cm. Seorang mahasiswa baru akan diukur tinggi badannya. Berapa peluang tinggi badannya berada dalam rentang $168 - 170$ cm.

$$P(Z \leq z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \, dt.$$



	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	0.5398	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	0.7704	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	0.9279	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
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

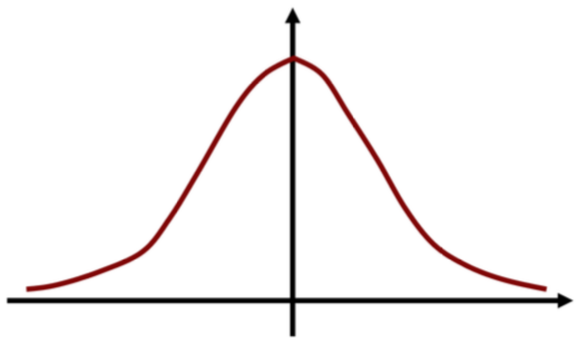
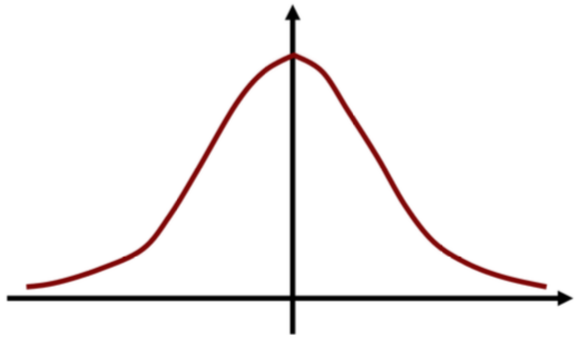
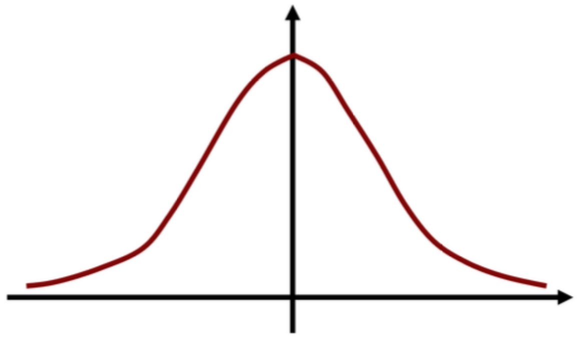
(a) Tinggi badan mahasiswa CIT didata dan ditemukan bahwa datanya berdistribusi normal dengan rata-rata 165 cm dan standar deviasi 18 cm. Seorang mahasiswa baru akan diukur tinggi badannya. Berapa peluang tinggi badannya berada dalam rentang $168 - 170$ cm.

https://colab.research.google.com/drive/1-FM1_nMWDqYezW4Z8fQw5ytkP4MO-5e6?usp=sharing.

(b) Sebuah meja kuliah akan didesain agar nyaman digunakan oleh semua orang dengan tinggi badan minimal H cm. Berapakah nilai H seharusnya ditentukan agar kita dapat dengan tingkat kepercayaan $72,5\%$ sehingga meja tersebut nyaman digunakan mahasiswa CIT?

Definisi. Misalkan $\mu, \sigma \in \mathbb{R}$ dan $\alpha \in [0, 1]$ dengan $X \sim \mathcal{N}(\mu, \sigma)$. Fungsi ____________________ atau *percent point function* merupakan nilai x terkecil sehingga $\mathbb{P}(X \leq x) = \alpha$ (tingkat kepercayaan).

Misalkan X berdistribusi normal, yakni $X \sim \mathcal{N}(\mu, \sigma)$. Maka, kita peroleh '*normal rules of thumb*':

Aturan	Deskripsi kejadian	Luas di bawah kurva	Nilai peluang
$\mu \pm \sigma$			68%
$\mu \pm 2\sigma$			95%
$\mu \pm 3\sigma$			99,7%

CALVIN INSTITUTE OF TECHNOLOGY

Untuk menyiapkan dataset yang baik untuk proyek Machine Learning dalam bidang Analisis Sentimen status Twitter, data mentah dari status user Twitter perlu dilakukan prapemrosesan.

- Banyaknya karakter dalam sebuah status Twitter diasumsikan mengikuti distribusi normal dengan rata-rata sebesar $117,53$ karakter dan standar deviasi $9,06$ karakter.
- Dalam algoritma prapemrosesan, suatu penyaring akan dipakai untuk menolak status Twitter yang terlalu pendek atau terlalu panjang, dengan target hanya membuang sekitar 5% dari dataset.

Gunakan '*normal rules of thumb*' untuk menentukan berapakah batas atas dan batas bawah banyak karakter yang seharusnya diijinkan oleh algoritma agar lolos prapemrosesan?

Distribusi geometri sering merepresentasikan waktu tunggu hingga percobaan Bernoulli sukses, sedangkan distribusi eksponensial menggambarkan waktu tunggu hingga kejadian dalam proses Poisson. Jika x adalah banyaknya kecelakaan pesawat dalam setahun dan $X \sim \text{Poi}(\lambda)$, maka waktu tunggu antar kecelakaan w mengikuti $W \sim \text{Exp}(\lambda)$, dengan λ sebagai laju kejadian per tahun.

iPhone X diketahui memiliki massa hidup rata-rata 6 tahun, mengikuti distribusi eksponensial. Berapa peluang sebuah iPhone milik Dr. Calvin dapat bertahan hingga delapan tahun?

Dalam suatu kantor Relasi Pelanggan, diketahui bahwa banyaknya telepon masuk mengenai keluhan pelanggan berdistribusi Poisson dengan laju 12 telepon per jam. Misalkan diketahui pada pk. 13.00 terdapat sebuah telepon keluhan pelanggan. Berapakah peluang telepon berikutnya masuk dalam rentang waktu antara pk. 13.03 hingga pk. 13.07?

CALVIN INSTITUTE OF TECHNOLOGY

2.3 Distribusi Peluang Kontinu Lainnya

Dalam statistik Bayesian, ____________________ merepresentasikan keyakinan awal tentang parameter sebelum data diamati, sedangkan ____________________ menghitung keyakinan yang diperbarui setelah mempertimbangkan data dengan Teorema Bayes. Hal ini dibahas lebih jauh dalam statistik lanjut.

Distribusi	Parameter	Notasi	*Support*	Karakteristik	Contoh	PDF: $f_X(x)$
Eksponensial						
Gamma						
Chi-Square						
t-student						
Beta						

Gambarkan sketsa fungsi kepadatan peluang $f_X(x)$ untuk setiap distribusi pada tabel di atas!