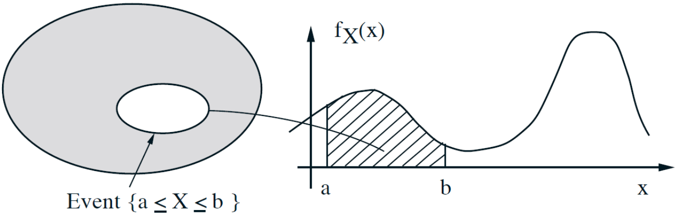
Continuous Random Variables

Definition

- A Continuous Random Variable is described by a **probability density function** f_X



$$\mathbf{P}(a \leq X \leq b) = \int_a^b f_X(x)\,dx$$

- Recall: for a continuous random variable X, $P(X = a) = 0$

Definition

- Probability that random variable X takes a value within a very small interval e.g. $X \in [x, x+\delta]$ for small δ:

$$\mathbf{P}(x \leq X \leq x+\delta) = \int_{x}^{x+\delta} f_X(s)\, ds \approx f_X(x) \cdot \delta$$

- Therefore, density at a particular point x, denoted by $f_X(x)$, can be approximated as:

$$f_X(x) \approx \frac{P(x \leq X \leq x+\delta)}{\delta}$$

- Density is not probability! It's probability per unit length.

- Density can be greater than 1, i.e. $f_X(x) > 1$.

- Density is always non-negative i.e. $f_X(x) \geq 0$.

Definition

- Let Ω be the entire sample space, i.e. $\Omega = (-\infty, +\infty)$, then:

$$1 = P(X \in \Omega) = P(X \in (-\infty, +\infty)) = \int_{-\infty}^{+\infty} f_X(x)\, dx$$

$$\Rightarrow \int_{-\infty}^{+\infty} f_X(x)\, dx = 1$$

- More generally,

$$\mathbf{P}(X \in B) = \int_B f_X(x)\, dx, \quad \text{for "nice" sets } B$$

For example,

$$P(X \in [1,2] \cup [3,4]) = \int_{[1,2]\cup[3,4]} f_X(x)\, dx = \int_1^2 f_X(x)\, dx + \int_3^4 f_X(x)\, dx$$

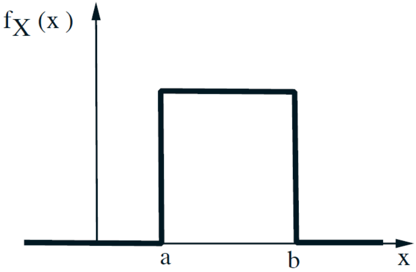
Mean and Variance

- Similar to discrete random variables, replace PMF with PDF, and sum with integral:

  - $\mathbf{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx$
  - $\mathbf{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx$
  - $\text{var}(X) = \sigma_X^2 = \int_{-\infty}^{\infty} (x - \mathbf{E}[X])^2 f_X(x)\,dx$

  and $\text{var}(X) = E[X^2] - E[X]^2$

Example



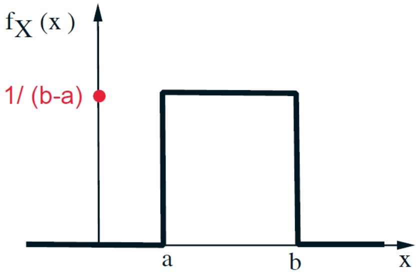
- What is height of the square? denote it by h.

$$\int_{-\infty}^{+\infty} f_X(x)\,dx = 1 \Rightarrow h \cdot (b-a) = 1 \Rightarrow h = \frac{1}{b-a}$$

Therefore,

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

Example

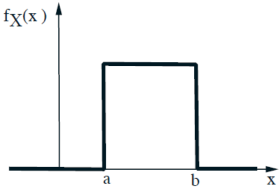


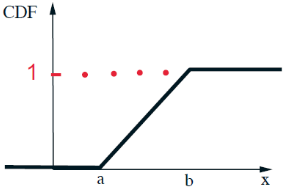
$$E[X] = \int_a^b x \cdot \frac{1}{b-a}\,dx = \frac{a+b}{2}$$

$$\mathrm{var}[X] = \int_a^b (x - \frac{a+b}{2})^2 \frac{1}{b-a}\,dx = \frac{(b-a)^2}{12}$$

Cumulative distribution function (CDF)

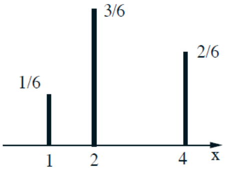
$$F_X(x) = \mathrm{P}(X \leq x) = \int_{-\infty}^{x} f_X(t)\,dt$$

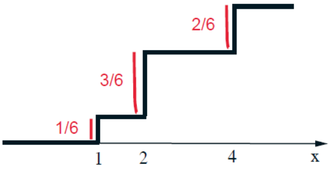




- Also for discrete r.v.'s:

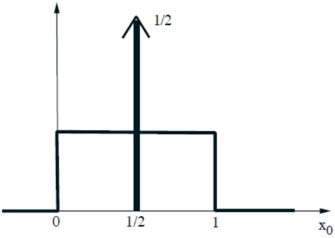
$$F_X(x) = \mathrm{P}(X \leq x) = \sum_{k \leq x} p_X(k)$$





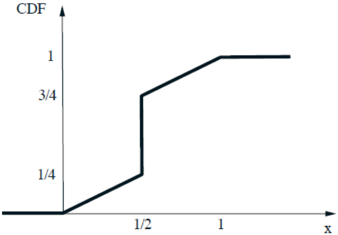
Mixed distributions

- Schematic drawing of a combination of a PDF and a PMF



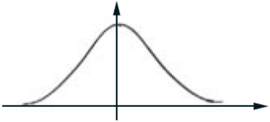
- The corresponding CDF:

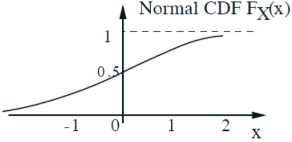
$$F_X(x) = \mathbf{P}(X \leq x)$$



Gaussian (normal) PDF

- Standard normal $N(0,1)$: $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$





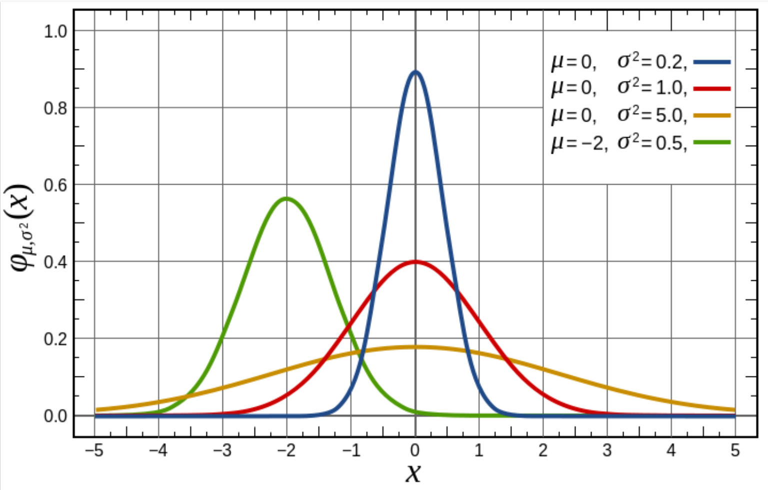
- $\mathbf{E}[X] = 0 \qquad \text{var}(X) = 1$

- General normal $N(\mu, \sigma^2)$:

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

- It turns out that:
  $\mathbf{E}[X] = \mu$ and $\text{Var}(X) = \sigma^2$.

Gaussian (normal) PDF



Gaussian (normal) PDF

- $X \sim N(\mu, \sigma^2)$ means X is a general normal random variable with $E[X] = \mu$, $\text{var}(X) = \sigma^2$.
- Let $Y = aX + b$ for constants a and b.
- Then by linearity of expectation $E[Y] = a\mu + b$.
- Also, $\text{var}(Y) = a^2\sigma^2$.
- Fact: Y will be a general normal random variable i.e. $Y \sim N(a\mu + b, a^2\sigma^2)$.

Gaussian (normal) PDF

- Let $X \sim N(\mu, \sigma^2)$ and $Y = \frac{X-\mu}{\sigma}$.
- What is distribution of Y?
- $Y = \frac{X-\mu}{\sigma} = \frac{X}{\sigma} + \frac{-\mu}{\sigma}$. Let $a = \frac{1}{\sigma}$ and $b = \frac{-\mu}{\sigma}$. Then $Y = aX + b$.
- Therefore, $Y \sim N(a\mu + b, a^2\sigma^2)$
- $a\mu + b = \frac{1}{\sigma}\mu + \frac{-\mu}{\sigma} = 0$.
- $a^2\sigma^2 = (\frac{1}{\sigma})^2\sigma^2 = 1$.
- Therefore, $Y \sim N(0, 1)$. So, Y is a standard normal random variable.

Conclusion:

$$\text{If } X \sim N(\mu, \sigma^2) \text{ then } \frac{X-\mu}{\sigma} \sim N(0, 1)$$

Example

- Fact: If $X \sim N(\mu, \sigma^2)$ then $\frac{X-\mu}{\sigma} \sim N(0,1)$.
- **Example 1**: Let $\mu = 2, \sigma = 4$. If $X \sim N(2,16)$ then $\frac{X-2}{4} \sim N(0,1)$

  So $\frac{X-2}{4}$ will be a standard normal random variable.
- **Example 2**: Let $\mu = 3, \sigma = 5$. If $X \sim N(3,25)$ then $\frac{X-3}{5} \sim N(0,1)$

  So $\frac{X-3}{5}$ will be a standard normal random variable.

Calculating normal probabilities

- There is no closed form available for CDF of a normal random variable but there are tables for standard normal random variable.

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817

Calculating normal probabilities

- How to find CDF of a general random variable using CDF of a nomral random variable?
- Suppose $X \sim N(2, 16)$ and we want to find CDF of X at point 3, i.e. $P(X \leq 3)$.
- We know If $X \sim N(2, 16)$ then $\frac{X-2}{4} \sim N(0, 1)$.
- $X \leq 3 \Leftrightarrow X - 2 \leq 3 - 2 \Leftrightarrow \frac{X-2}{4} \leq \frac{3-2}{4} = 0.25$
- Therefore, $X \leq 3 \Leftrightarrow \frac{X-2}{4} \leq 0.25$.
- It means that "$X \leq 3$" and "$\frac{X-2}{4} \leq 0.25$" are the same event.
- Therefore, $P(X \leq 3) = P(\frac{X-2}{4} \leq 0.25)$.
- But $\frac{X-2}{4}$ is a standard normal, so $P(\frac{X-2}{4} \leq 0.25)$ is CDF of a standard normal at point 0.25.
- Based on the table, CDF(0.25)=0.5987. So $P(X \leq 3) = 0.5987$.