MATH 180A (Lecture A00)

mathweb.ucsd.edu/~ynemish/teaching/180a

Today: Gaussian distribution

Next: ASV 4.1

Week 6:

- Homework 4 due Friday, February 17

Variance

Definition The variance of a random variable X is

$$\mathrm{Var}(X) = E\left(\left(X - E(X)\right)^2\right)$$

Proposition. Let X be a random variable. Then

$$\mathrm{Var}(X) = E(X^2) - \left(E(X)\right)^2$$

- The square root of the variance is called standard deviation $\sigma(X) = \sqrt{\mathrm{Var}(X)}$

Variance

Variance is a measure of how "spread out from the mean" the distribution is.

Proposition Let X be a random variable with finite expectation $E(X)=\mu$. Then

$$\operatorname{Var}(X)=0 \quad \text{iff} \quad P(X=\mu)=1$$

Proof $(\Leftarrow)$ Exercise

$(\Rightarrow)$ (Assume X is discrete).

$$0=\operatorname{Var}(X)=\sum_t \underbrace{(t-\mu)^2}_{\geq 0}\underbrace{P(X=t)}_{\geq 0} \Rightarrow \text{For all } t,\ (t-\mu)^2 P(X=t)=0$$

For all t, either $(t-\mu)^2=0$ or $P(X=t)=0$ so if $t\neq\mu$, then $P(X=t)=0$

therefore, $P(X=\mu)=1$

Expectation and variance of $aX+b$

Let X be a random variable, and let $a, b \in \mathbb{R}$. Then

(i) $E(aX+b) = aE(X)+b$

(ii) $Var(aX+b) = a^2 Var(X)$ if $E(X)$ and $Var(X)$ exist

Proof (i) $\leftarrow$ homework

(ii)
$$
\begin{aligned}
Var(aX+b) &= E\left(\left(aX+b-E(aX+b)\right)^2\right) \\
&\overset{(i)}{=} E\left(\left(aX+\cancel{b}-aE(X)-\cancel{b}\right)^2\right) \\
&= E\left(a^2(X-E(X))^2\right) \\
&\overset{(i)}{=} a^2 E\left((X-E(X))^2\right) \\
&= a^2 Var(X)
\end{aligned}
$$

Variance of geometric distribution

Let $X \sim \text{Geom}(p)$. We know that $E(X) = \frac{1}{p}$

$$E(X^2) = \sum_{k=1}^{\infty} k^2 P(X=k) = \sum_{k=1}^{\infty} k^2 \cdot p(1-p)^{k-1} = \sum_{k=1}^{\infty} \left((k-1)+1\right)^2 p(1-p)^{k-1}$$

$$= \sum_{k=1}^{\infty} (k-1)^2 p(1-p)^{k-1} + 2 \sum_{k=1}^{\infty} (k-1) p(1-p)^{k-1} + \underbrace{\sum_{k=1}^{\infty} p(1-p)^{k-1}}_{1}$$

$\ell = k-1$ $$= \sum_{\ell=1}^{\infty} \ell^2 p(1-p)^{\ell} + 2 \sum_{\ell=1}^{\infty} \ell \, p(1-p)^{\ell} + 1$$

$$= (1-p) \sum_{\ell=1}^{\infty} \ell^2 p(1-p)^{\ell-1} + 2 \cdot (1-p) \sum_{\ell=1}^{\infty} \ell \, p(1-p)^{\ell-1} + 1$$

$$= (1-p) E(X^2) + 2 \cdot (1-p) \cdot E(X) + 1$$

$$= (1-p) E(X^2) + 2 \cdot (1-p) \cdot \frac{1}{p} + 1 \Rightarrow E(X^2) = (1-p) E(X^2) + \frac{2}{p} - 1$$

$$p E(X^2) = \frac{2-p}{p}, \quad E(X^2) = \frac{2-p}{p^2} \Rightarrow \text{Var}(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

Random variables. Summary

Discrete	Continuous
Finite/countable set of possible values, $\sum_t P(X=t)=1$	Uncountable set of possible values, $\forall t \in \mathbb{R}$ $P(X=t)=0$
PMF: $p_X(t)=P(X=t)$ $P(X\in B)=\sum_{t\in B} p_X(t)$	PDF: $f_X:\mathbb{R}\to\mathbb{R}$ $P(X\in B)=\int_B f_X(t)\,dt$
CDF F_X is a step function	CDF F_X is a continuous function
Expectation: $E(X)=\sum_t t\,P(X=t)$	Expectation: $E(X)=\int_{\mathbb{R}} t\,f_X(t)\,dt$
$E(g(X))=\sum_t g(t)\,P(X=t)$	$E(g(X))=\int_{\mathbb{R}} g(t)\,f_X(t)\,dt$
Relation between CDF and PMF: magnitude of jump of F_X at t is $P(X=t)$	Relation between CDF and PDF: $f_X(t)=F_X'(t)$ on the intervals where F_X is differentiable

Random variables. Summary

F_X is (i) nondecreasing, (ii) right-continuous

$$\text{(iii)} \lim_{t \to -\infty} F_X(t) = 0, \quad \lim_{t \to +\infty} F_X(t) = 1$$

Variance: $\operatorname{Var}(X) = E\left((X - E(X))^2 \right) = E(X^2) - \left(E(X)\right)^2$

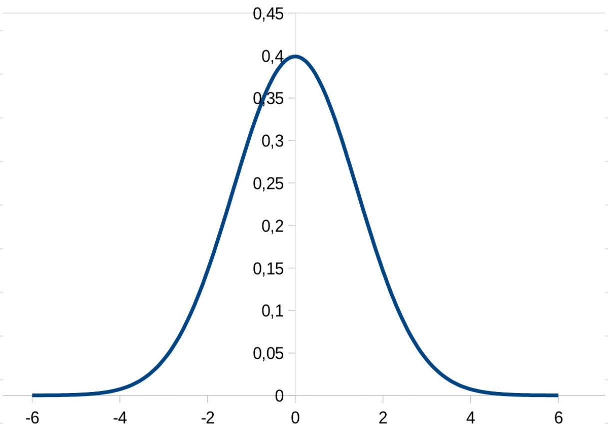
$$E(aX+b) = aE(X) + b, \qquad \operatorname{Var}(aX+b) = a^2 \operatorname{Var}(X)$$

Gaussian (normal) distribution

Gaussian (normal) distribution

Def Random variable Z has the standard normal (Gaussian) distribution if the PDF of Z is given by

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$



Notation: $Z \sim \mathcal{N}(0,1)$

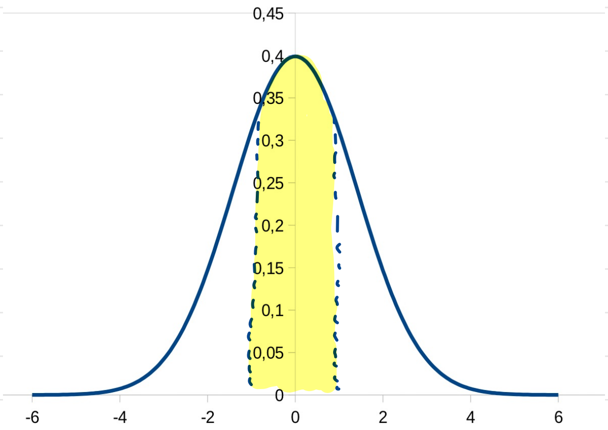
Is $\varphi(t)$ indeed a PDF?

- $\varphi(t) \geq 0$ ✓
- $\int_{-\infty}^{+\infty} \varphi(t)\,dt = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt \overset{?}{=} 1$ ✓

$$I^2 = \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}dt \int_{-\infty}^{+\infty} e^{-\frac{s^2}{2}}ds = \iint_{\mathbb{R}^2} e^{-\frac{t^2+s^2}{2}}\,ds\,dt \overset{\text{polar coord}}{=} \int_0^{2\pi}\int_0^{+\infty} e^{-\frac{r^2}{2}}\, r\,dr\,d\theta = 2\pi \int_0^{+\infty} e^{-\frac{r^2}{2}} d\left(\frac{r^2}{2}\right) = 2\pi$$

CDF of $N(0,1)$

Suppose $X \sim N(0,1)$. What is $P(|X| \leq 1)$?



$$P(-1 \leq X \leq 1) = \int_{-1}^{1} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-\frac{t^2}{2}} dt$$

Cannot use the polar coordinate trick.

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$ - CDF of $X \sim N(0,1)$

$\Phi(1) - \Phi(-1)$

- no simple explicit formula
- table of values of $\Phi(x)$ (for $x \geq 0$)

Normal table of values (Appendix E in textbook)

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	0.5398	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

This table gives $P(Z \le z)$ where $Z \sim N(0,1)$, $z = x_i + y_j$

<u>Example</u> $\Phi(0.91) = P(Z \le 0.91) = P(Z \le 0.9 + 0.01) \approx 0.8186$

<u>Fact</u>:

$P(Z > 0.24) =$

$P(-0.28 < Z < 0.59) =$