LECTURE 8

- **Readings:** Sections 3.1–3.3

Lecture outline

- Review and summary
- Probability density functions
- Cumulative distribution functions
- Normal random variables

Review: Variance in the hat problem

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

$$X^2 = \sum_i X_i^2 + \sum_{i,j:i \neq j} X_i X_j$$

- $\mathbf{E}[X_i^2] = \mathbf{P}(X_i = 0) = \frac{1}{n}$
- For $i \neq j$,

$$\begin{aligned} \mathbf{E}[X_i X_j] &= \mathbf{P}(X_i = 1,\, X_j = 1) \\ &= \underbrace{\mathbf{P}(X_i = 1)}_{1/n}\, \underbrace{\mathbf{P}(X_j = 1 \mid X_i = 1)}_{1/(n-1)} \end{aligned}$$

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

$$= \left(n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n(n-1)}\right) - (1)^2 = 1$$

Summary

- If you know what every notation means . . .

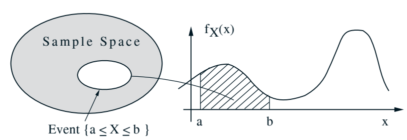
A, B: events; X, Y: random variables

$\mathbf{P}(B)$	$\mathbf{P}(B \mid A)$	$p_{X,Y}(x, y)$
$p_X(x)$	$p_{X\mid A}(x)$	$p_{X\mid Y}(x \mid y)$
$\mathbf{E}[X]$	$\mathbf{E}[X \mid A]$	$\mathbf{E}[g(X)]$
$\text{var}(X)$	$\text{var}(X \mid A)$	σ_X

- Independence
- Special random variables:
  discrete uniform, Bernoulli, binomial, geometric, Poisson
- Quiz 1 has multiple-choice and detailed-answer questions

Continuous random variables

- Recall: a random variable assigns a number to each outcome
- **Continuous** random variable X is described by **probability density function** f_X



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

Probability density functions

- "Density" of probability: When $\delta > 0$ is very small

$$\mathbf{P}(X \in [x, x+\delta]) = \int_x^{x+\delta} f_X(x)\, dx \approx f_X(x) \cdot \delta$$

$$f_X(x) \approx \frac{\mathbf{P}(X \in [x, x+\delta])}{\delta}$$

- Properties:
  - $f_X(x) \geq 0$ for every x
  - When B is a countable union of intervals,

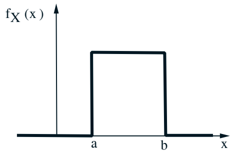
$$\mathbf{P}(X \in B) = \int_B f_X(x)\, dx$$

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

Means and variances

- $\mathbf{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx$
- $\mathbf{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$
  - $\mathbf{E}[aX + b] = a\mathbf{E}[X] + b$ for scalars a and b
- $\text{var}(X) = \sigma_X^2 = \int_{-\infty}^{\infty} (x - \mathbf{E}[X])^2 f_X(x)\, dx$
  - $\text{var}(aX + b) = a^2\text{var}(X)$ for scalars a and b

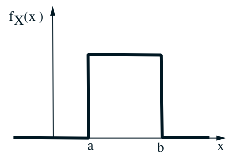
Continuous uniform random variable

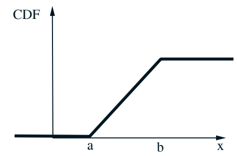


- $f_X(x) \ =$ for $a \leq x \leq b$
- $\mathbf{E}[X] \ =$
- var$(X) \ =$

Cumulative distribution function (CDF)

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





CDFs of continuous random variables

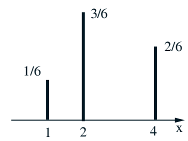
$$F_X(x) \ = \ \mathbf{P}(X \leq x) \ = \ \int_{-\infty}^{x} f_X(t)\, dt$$

- $F_X(x)$ is monotonically nondecreasing:

  if $x \leq y$, then $F_X(x) \leq F_X(y)$
- $F_X(x)$ tends to 0 as $x \to -\infty$, and to 1 as $x \to \infty$
- $F_X(x)$ is a continuous function of x
- $f_X(x) = \dfrac{d}{dx} F_X(x)$ where the derivative exists

CDF for discrete random variables

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



CDFs of discrete random variables

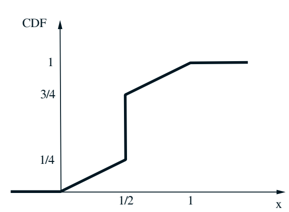
$$F_X(x) \ = \ \mathbf{P}(X \leq x) \ = \ \sum_{k \leq x} p_X(k)$$

- $F_X(x)$ is monotonically nondecreasing:

  if $x \leq y$, then $F_X(x) \leq F_X(y)$
- $F_X(x)$ tends to 0 as $x \to -\infty$, and to 1 as $x \to \infty$
- $F_X(x)$ is a piecewise constant function of x
- $p_X(x)$ equals size of jump of F_X at x

Mixed random variables

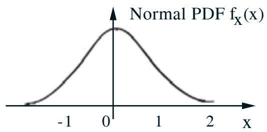
- Random variable can be neither discrete nor continuous
- CDF $F_X(x) = \mathbf{P}(X \leq x)$ still well defined

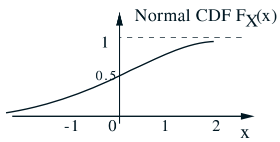


- $F_X(x)$ is monotonically nondecreasing
- $F_X(x)$ tends to 0 as $x \to -\infty$, and to 1 as $x \to \infty$

Normal (Gaussian) PDF

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





- $\mathbf{E}[X] =$ $\text{var}(X) = 1$
- General normal $N(\mu, \sigma^2)$:

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

Calculating normal probabilities

- No closed form available for normal CDF
- Define a new function Φ as the standard normal CDF
  - Implemented in many software packages
  - Can be found in tables
- Fact (Section 4.1): If X is normal, then $aX + b$ is normal
- If $X \sim N(\mu, \sigma^2)$, then $\dfrac{X-\mu}{\sigma} \sim N(\qquad)$
- If $X \sim N(2, 16)$:

$$\mathbf{P}(X \le 3) \;=\; \mathbf{P}\left(\frac{X-2}{4} \le \frac{3-2}{4}\right) \;=\; \Phi(0.25)$$

Standard normal table

	.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