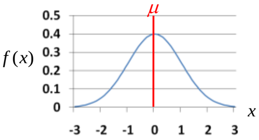
Normal Random Variable

- X is a **Normal Random Variable**: $X \sim N(\mu, \sigma^2)$
  - Probability Density Function (PDF):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad \text{where} \quad -\infty < x < \infty$$

  - $E[X] = \mu$
  - $Var(X) = \sigma^2$



  - Also called "Gaussian"
  - Note: $f(x)$ is symmetric about μ
  - Common for natural phenomena: heights, weights, etc.
  - Often results from the sum of multiple variables

Carl Friedrich Gauss

- Carl Friedrich Gauss (1777-1855) was a remarkably influential German mathematician

- Started doing groundbreaking math as teenager
  - Did not invent Normal distribution, but popularized it
- He looked more like Martin Sheen
  - Who is, of course, Charlie Sheen's father

Properties of Normal Random Variable

- Let $X \sim N(\mu, \sigma^2)$
- Let $Y = aX + b$
  - $Y \sim N(a\mu + b, a^2\sigma^2)$
  - $E[Y] = E[aX + b] = aE[X] + b = a\mu + b$
  - $Var(Y) = Var(aX + b) = a^2 Var(X) = a^2\sigma^2$

$$F_Y(x) = P(Y \le x) = P(aX + b \le x) = P(X \le \tfrac{x-b}{a}) = F_X(\tfrac{x-b}{a})$$

Differentiating $F_Y(x)$ w.r.t. x, yields $f_Y(x)$, the PDF for y:

$$f_Y(x) = \frac{d}{dx} F_Y(x) = \frac{d}{dx} F_X(\tfrac{x-b}{a}) = \tfrac{1}{a} f_X(\tfrac{x-b}{a})$$

- Special case: $Z = (X - \mu)/\sigma$ $\quad (a = 1/\sigma,\ b = -\mu/\sigma)$
  - $Z \sim N(a\mu + b, a^2\sigma^2) = N(\mu/\sigma - \mu/\sigma, (1/\sigma)^2\sigma^2) = N(0, 1)$

Standard (Unit) Normal Random Variable

- Z is a **Standard (or Unit) Normal RV**: $Z \sim N(0, 1)$
  - $E[Z] = \mu = 0$ $\quad Var(Z) = \sigma^2 = 1$ $\quad SD(Z) = \sigma = 1$
  - CDF of Z, $F_Z(z)$ does not have closed form
  - We denote $F_Z(z)$ as $\Phi(z)$: "phi of z"

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

  - By symmetry: $\Phi(-z) = P(Z \le -z) = P(Z \ge z) = 1 - \Phi(z)$
- Use Z to compute $X \sim N(\mu, \sigma^2)$, where $\sigma > 0$

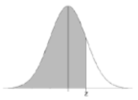
$$F_X(x) = P(X \le x) = P(\tfrac{X-\mu}{\sigma} \le \tfrac{x-\mu}{\sigma}) = P(Z \le \tfrac{x-\mu}{\sigma}) = \Phi(\tfrac{x-\mu}{\sigma})$$

  - Table of $\Phi(z)$ values in textbook, p. 201 and handout

Using Table of Φ(z) Values

Standard Normal Cumulative Probability Table

$\Phi(0.54) = 0.7054$



Cumulative probabilities for POSITIVE z-values are shown in the following table:

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
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

Get Your Gaussian On

- $X \sim N(3, 16)$ $\quad \mu = 3$ $\quad \sigma^2 = 16$ $\quad \sigma = 4$
  - What is $P(X > 0)$?

$$P(X > 0) = P(\tfrac{X-3}{4} > \tfrac{0-3}{4}) = P(Z > -\tfrac{3}{4}) = 1 - P(Z \le -\tfrac{3}{4})$$

$$1 - \Phi(-\tfrac{3}{4}) = \Phi(\tfrac{3}{4}) = 0.7734$$

  - What is $P(2 < X < 5)$?

$$P(2 < X < 5) = P(\tfrac{2-3}{4} < \tfrac{X-3}{4} < \tfrac{5-3}{4}) = P(-\tfrac{1}{4} < Z < \tfrac{2}{4})$$

$$\Phi(\tfrac{2}{4}) - \Phi(-\tfrac{1}{4}) = \Phi(\tfrac{1}{2}) - (1 - \Phi(\tfrac{1}{4})) = 0.6915 - (1 - 0.5987) = 0.2902$$

  - What is $P(|X - 3| > 6)$?

$$P(X < -3) + P(X > 9) = P(Z < \tfrac{-3-3}{4}) + P(Z > \tfrac{9-3}{4})$$

$$\Phi(-\tfrac{3}{2}) + (1 - \Phi(\tfrac{3}{2})) = 2(1 - \Phi(\tfrac{3}{2})) = 2(1 - 0.9332) = 0.1336$$

Noisy Wires

- Send voltage of 2 or -2 on wire (to denote 1 or 0)
  - X = voltage sent
  - R = voltage received = X + Y, where noise Y ~ N(0, 1)
  - Decode R: if (R ≥ 0.5) then 1, else 0
  - What is P(error after decoding | original bit = 1)?

$$P(2+Y<0.5)=P(Y<-1.5)=\Phi(-1.5)=1-\Phi(1.5)\approx 0.0668$$

  - What is P(error after decoding | original bit = 0)?

$$P(-2+Y\geq 0.5)=P(Y\geq 2.5)=1-\Phi(2.5)\approx 0.0062$$

Normal Approximation to Binomial

- X ~ Bin(n, p)
  - E[X] = np Var(X) = np(1 – p)
  - Poisson approx. good: n large (> 20), p small (< 0.05)
  - For large n: X ≈ Y ~ N(E[X], Var(X)) = N(np, np(1 – p))
  - Normal approx. good : Var(X) = np(1 – p) ≥ 10

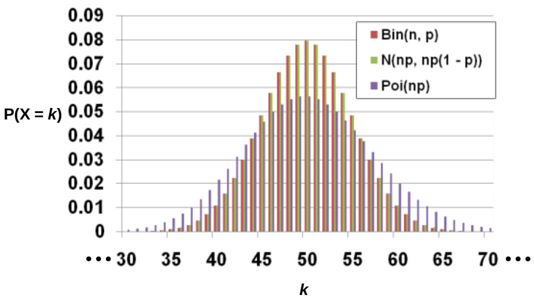
$$P(X=k)\approx P\left(k-\tfrac{1}{2}<Y<k+\tfrac{1}{2}\right)=\Phi\left(\frac{k-np+0.5}{\sqrt{np(1-p)}}\right)-\Phi\left(\frac{k-np-0.5}{\sqrt{np(1-p)}}\right)$$

"Continuity correction"

  - DeMoivre-Laplace Limit Theorem:
    - S_n: number of successes (with prob. p) in n independent trials

$$P\left(a\leq\frac{S_n-np}{\sqrt{np(1-p)}}\leq b\right)\xrightarrow{n\to\infty}\Phi(b)-\Phi(a)$$

Comparison when n = 100, p = 0.5



Faulty Endorsements

- 100 people placed on special diet
  - X = # people on diet whose cholesterol decreases
  - Doctor will endorse diet if X ≥ 65
  - What is P(doctor endorses diet | diet has no effect)?
  - X ~ Bin(100, 0.5)

$$np = 50 \quad np(1-p) = 25 \quad \sqrt{np(1-p)} = 5$$

  - Use Normal approximation: Y ~ N(50, 25)

$$P(X\geq 65)\approx P(Y>64.5)$$

$$P(Y\geq 64.5)=P\left(\tfrac{Y-50}{5}>\tfrac{64.5-50}{5}\right)=1-\Phi(2.9)\approx 0.0019$$

  - Using Binomial:

$$P(X\geq 65)\approx 0.0018$$

Stanford Admissions

- Stanford accepts 2480 students
  - Each accepted student has 68% chance of attending
  - X = # students who will attend. X ~ Bin(2480, 0.68)
  - What is P(X > 1745)?

$$np = 1686.4 \quad np(1-p)\approx 539.65 \quad \sqrt{np(1-p)}\approx 23.23$$

  - Use Normal approximation: Y ~ N(1686.4, 539.65)

$$P(X>1745)\approx P(Y\geq 1745.5)$$

$$P(Y\geq 1745.5)=P\left(\tfrac{Y-1686.4}{23.23}>\tfrac{1745.5-1686.4}{23.23}\right)=1-\Phi(2.54)\approx 0.0055$$

  - Using Binomial:

$$P(X>1745)\approx 0.0053$$

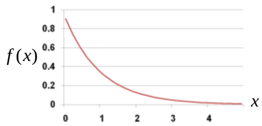
And now for something (mostly) completely different…

Exponential Random Variable

- X is an **Exponential RV**: $X \sim \text{Exp}(\lambda)$ Rate: $\lambda > 0$
  - Probability Density Function (PDF):

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad \text{where } -\infty < x < \infty$$

  - $E[X] = \dfrac{1}{\lambda}$
  - $Var(X) = \dfrac{1}{\lambda^2}$



  - Cumulative distribution function (CDF), $F(X) = P(X \leq x)$:

$$F(x) = 1 - e^{-\lambda x} \quad \text{where } x \geq 0$$

  - Represents time until some event
    - Earthquake, request to web server, end cell phone contract, etc.

Exponential is "Memoryless"

- X = time until some event occurs
  - X ~ Exp(λ)
  - What is P(X > s + t | X > s)?

$$P(X > s+t \mid X > s) = \frac{P(X > s+t \text{ and } X > s)}{P(X > s)} = \frac{P(X > s+t)}{P(X > s)}$$

$$\frac{P(X > s+t)}{P(X > s)} = \frac{1 - F(s+t)}{1 - F(s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = 1 - F(t) = P(X > t)$$

$$\text{So, } P(X > s+t \mid X > s) = P(X > t)$$

  - After initial period of time *s*, P(X > t | •) for waiting another *t* units of time until event is same as at start
  - "Memoryless" = no impact from preceding period *s*

Visits to Web Site

- Say a visitor to your web leaves after X minutes
  - On average, visitors leave site after 5 minutes
  - Assume length of stay is Exponentially distributed
  - X ~ Exp(λ = 1/5), since E[X] = 1/λ = 5
  - What is P(X > 10)?

$$P(X > 10) = 1 - F(10) = 1 - (1 - e^{-\lambda 10}) = e^{-2} \approx 0.1353$$

  - What is P(10 < X < 20)?

$$P(10 < X < 20) = F(20) - F(10) = (1 - e^{-4}) - (1 - e^{-2}) \approx 0.1170$$

Replacing Your Laptop

- X = # hours of use until your laptop dies
  - On average, laptops die after 5000 hours of use
  - X ~ Exp(λ = 1/5000), since E[X] = 1/λ = 5000
  - You use your laptop 5 hours/day.
  - What is P(your laptop lasts 4 years)?
  - That is: P(X > (5)(365)(4) = 7300)

$$P(X > 7300) = 1 - F(7300) = 1 - (1 - e^{-7300/5000}) = e^{-1.46} \approx 0.2322$$

  - Better plan ahead... especially if you are coterming:

$$P(X > 9125) = 1 - F(9125) = e^{-1.825} \approx 0.1612 \quad \text{(5 year plan)}$$

$$P(X > 10950) = 1 - F(10950) = e^{-2.19} \approx 0.1119 \quad \text{(6 year plan)}$$

A Little Calculus Review

- Product rule for derivatives:

$$d(u \cdot v) = du \cdot v + u \cdot dv$$

- Derivative and integral of exponential:

$$\frac{d(e^u)}{dx} = e^u \frac{du}{dx} \qquad \int e^u du = e^u$$

- Integration by parts:

$$\int d(u \cdot v) = u \cdot v = \int v \cdot du + \int u \cdot dv$$

$$\int u \cdot dv = u \cdot v - \int v \cdot du$$

And Now, Some Calculus Practice

- Compute *n*-th moment of Exponential distribution

$$E[X^n] = \int_0^\infty x^n \lambda e^{-\lambda x} dx$$

  - Step 1: don't panic, think happy thoughts, recall...
  - Step 2: find *u* and *v* (and *du* and *dv*):

$$u = x^n \qquad v = -e^{-\lambda x}$$

$$du = nx^{n-1}dx \qquad dv = \lambda e^{-\lambda x}dx$$

  - Step 3: substitute (a.k.a. "plug and chug")

$$\int u \cdot dv = \int x^n \cdot \lambda e^{-\lambda x} dx = u \cdot v - \int v \cdot du = -x^n e^{-\lambda x} + \int nx^{n-1} e^{-\lambda x} dx$$

$$E[X^n] = -x^n e^{-\lambda x}\Big|_0^\infty + \int nx^{n-1} e^{-\lambda x} dx = 0 + \tfrac{n}{\lambda} \int x^{n-1} \lambda e^{-\lambda x} dx = \tfrac{n}{\lambda} E[X^{n-1}]$$

$$\text{Base case}: E[X^0] = E[1] = 1, \quad \text{so } E[X] = \tfrac{1}{\lambda}, E[X^2] = \tfrac{2}{\lambda}\tfrac{1}{\lambda} = \tfrac{2}{\lambda^2}, \ldots$$