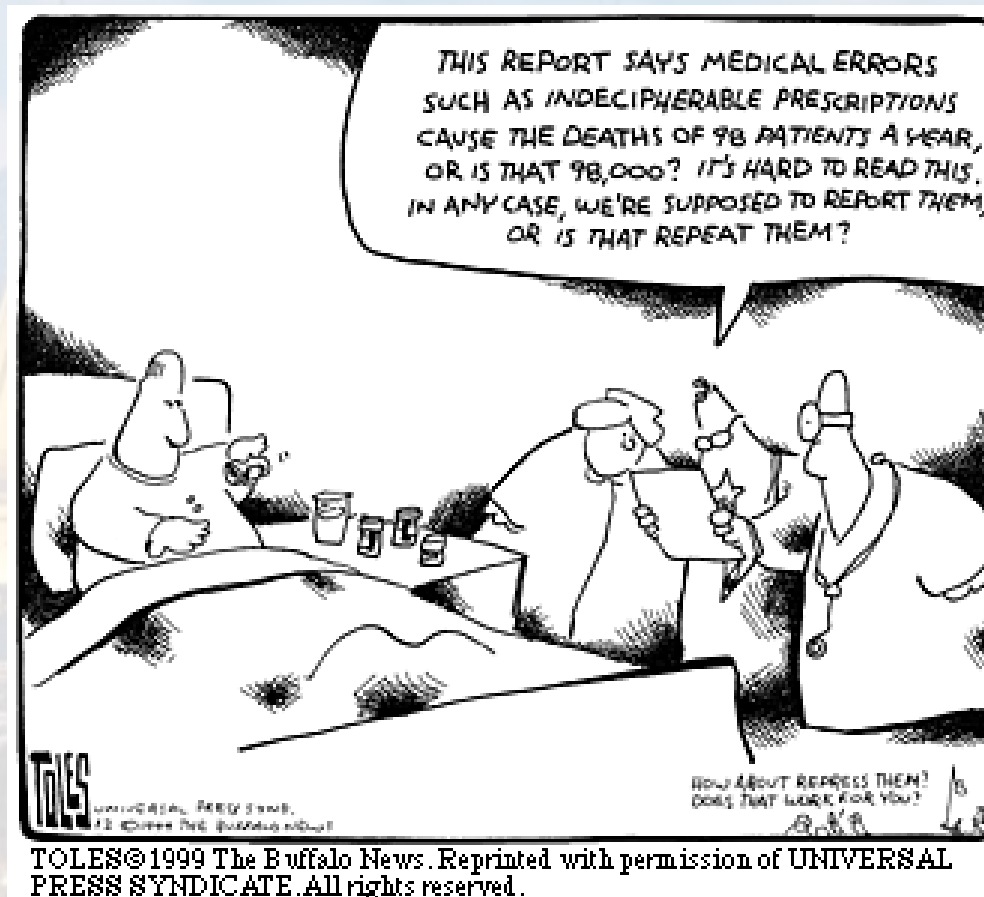


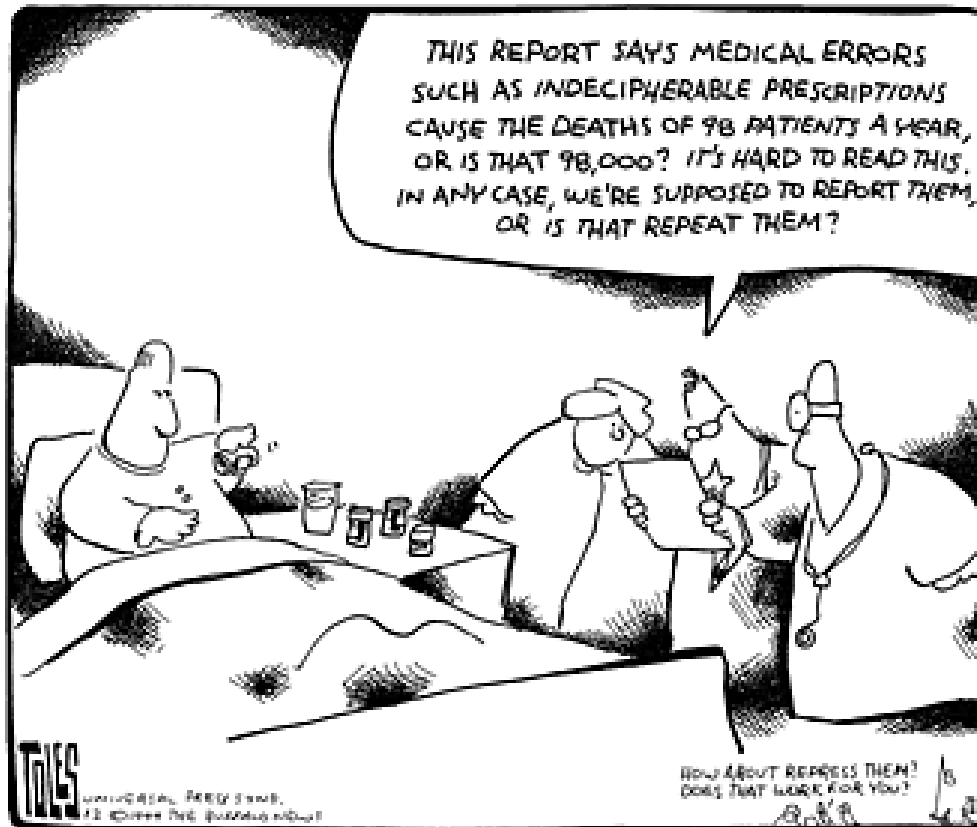
M. Hohle:

Physics 77: Introduction to Computational Techniques in Physics



syllabus:

- Introduction to Unix & Python (week 1 - 2)
- Functions, Loops, Lists and Arrays (week 3 - 4)
- Visualization (week 5)
- Parsing, Data Processing and File I/O (week 6)
- **Statistics and Probability, Interpreting Measurements (week 7 - 8)**
- **Random Numbers, Simulation (week 9)**
- Numerical Integration and Differentiation (week 10)
- Root Finding, Interpolation (week 11)
- Systems of Linear Equations (week 12)
- Ordinary Differential Equations (week 13)
- Fourier Transformation and Signal Processing (week 14)
- Capstone Project Presentations (week 15)



TOLES©1999 The Buffalo News. Reprinted with permission of UNIVERSAL PRESS SYNDICATE. All rights reserved.

Outline:

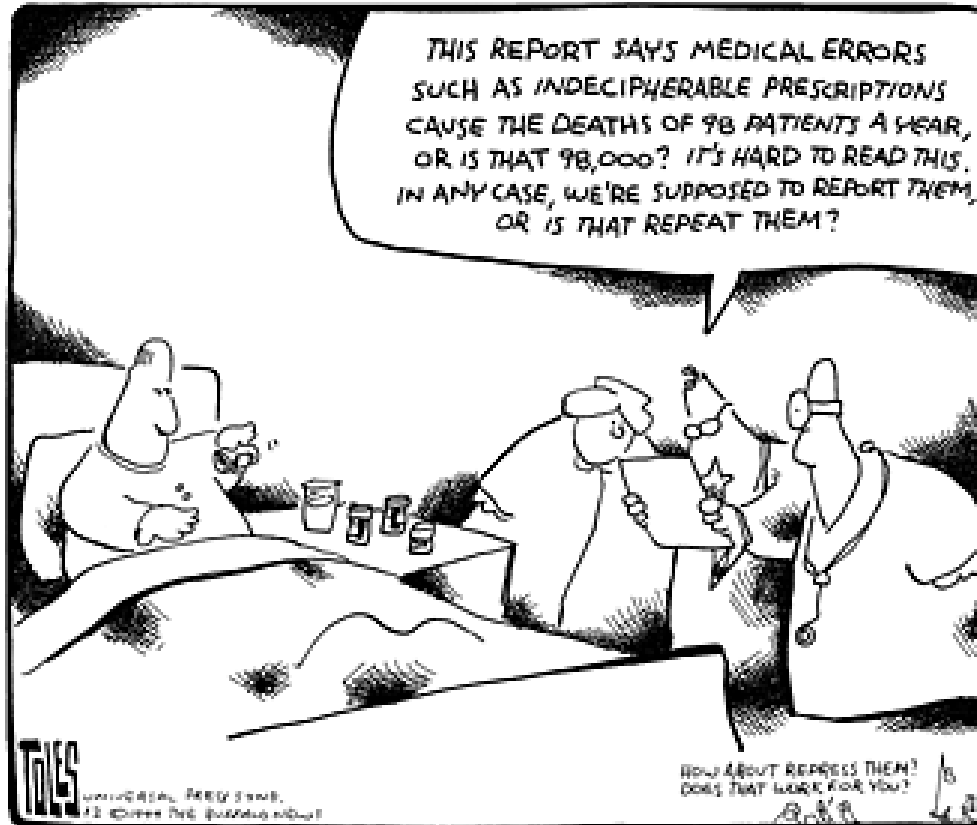
Basics

Most Common PDFs

- uniform
- binomial
- Poissonian
- Normal/Gaussian

Error Estimation

Bayesian Statistics



TOLES©1999 The Buffalo News. Reprinted with permission of UNIVERSAL PRESS SYNDICATE. All rights reserved.

Outline:

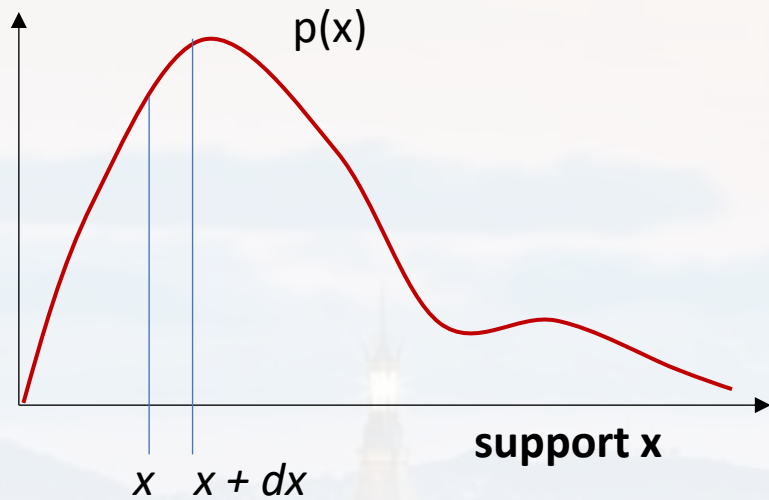
Basics

Most Common PDFs

- uniform
- binomial
- Poissonian
- Normal/Gaussian

Error Estimation

Bayesian Statistics



probability **d**ensity **f**unction

$[a \leq x \leq b]$

$$p(x) dx \quad \int_a^b p(x) dx = 1$$

Cumulative **d**ensity **f**unction

$$P(x) = \int_a^x p(x) dx$$

mean

$$\mu = E(x) = \int x p(x) dx$$

What is the mean of a fair, six sided die?

variance

$$\sigma^2 = \text{var}(x) = \int (x - \mu)^2 p(x) dx$$

median m

$$\int_a^m p(x) dx = \frac{1}{2}$$



mean

$$\mu = E(x) = \int x p(x) dx$$

variance

$$\sigma^2 = \text{var}(x) = \int (x - \mu)^2 p(x) dx$$

$$\text{var}(x) = \int (x - \mu)^2 p(x) dx = E([x - \mu]^2)$$

$$= E(x^2 - 2x\mu + \mu^2)$$

$$= E(x^2) - 2\mu E(x) + \mu^2 E(1)$$

$$= E(x^2) - 2\mu E(x) + \mu^2$$

$$\sigma^2 = E(x^2) - E(x)^2$$

the mean is a linear operator

$$\int_a^b p(x) dx = 1$$

$$\mu = E(x)$$



a, b = const

mean

$$\mu = E(x) = \int x p(x) dx$$

variance

$$\sigma^2 = \text{var}(x) = \int (x - \mu)^2 p(x) dx$$

$$\text{var}([a x_1 + b x_2]) = E([a x_1 + b x_2]^2) - E(a x_1 + b x_2)^2$$

$$\sigma^2 = E(x^2) - E(x)^2$$

$$= E(a^2 x_1^2 + 2ab x_1 x_2 + b^2 x_2^2) - E(a x_1 + b x_2)^2$$

$$= a^2 E(x_1^2) + 2ab E(x_1 x_2) + b^2 E(x_2^2) - E(a x_1 + b x_2)^2$$

$$= a^2 E(x_1^2) + 2ab E(x_1 x_2) + b^2 E(x_2^2) - [aE(x_1) + bE(x_2)]^2$$

$$= \boxed{a^2 E(x_1^2)} + 2ab E(x_1 x_2) + \boxed{b^2 E(x_2^2)} - \boxed{a^2 E(x_1)^2} - \boxed{b^2 E(x_2)^2} - 2ab E(x_1) E(x_2)$$

$$a^2 \text{var}(x_1)$$

$$b^2 \text{var}(x_2)$$



a, b = const

mean

$$\mu = E(x) = \int x p(x) dx$$

variance

$$\sigma^2 = \text{var}(x) = \int (x - \mu)^2 p(x) dx$$

$$\text{var}([a x_1 + b x_2]) = E([a x_1 + b x_2]^2) - E(a x_1 + b x_2)^2$$

$$\sigma^2 = E(x^2) - E(x)^2$$

$$= \boxed{a^2 E(x_1^2)} + \boxed{2ab E(x_1 x_2)} + \boxed{b^2 E(x_2^2)} - \boxed{a^2 E(x_1)^2} - \boxed{b^2 E(x_2)^2} - \boxed{2ab E(x_1) E(x_2)}$$

$a^2 \text{var}(x_1)$
 $b^2 \text{var}(x_2)$
 $2ab \text{cov}(x_1, x_2)$

$$= a^2 \text{var}(x_1) + b^2 \text{var}(x_2) + 2ab \text{cov}(x_1, x_2)$$

$$\text{cov}(x_1, x_2) = \text{cov}(x_2, x_1) = E(x_1 x_2) - E(x_1) E(x_2)$$

covariance

$$ab = c$$

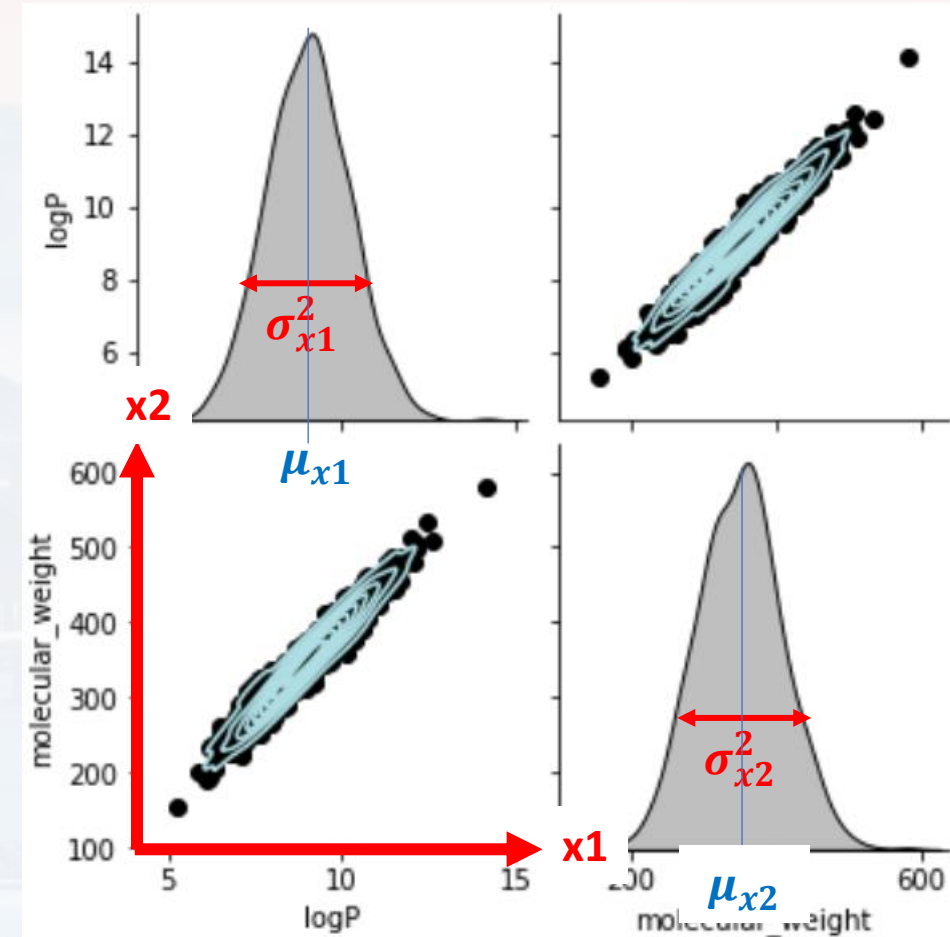
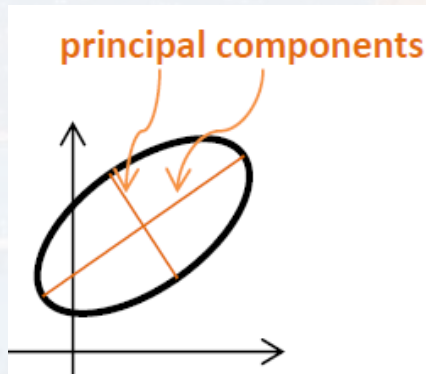


$$\text{const} = a^2 \text{var}(x_1) + b^2 \text{var}(x_2) + 2c \text{cov}(x_1, x_2)$$

$$\text{const} = \begin{pmatrix} x_1 - \mu_{x1} \\ x_2 - \mu_{x2} \end{pmatrix}^T \begin{pmatrix} a & c_{12} \\ c_{21} & b \end{pmatrix} \begin{pmatrix} x_1 - \mu_{x1} \\ x_2 - \mu_{x2} \end{pmatrix}$$

covariance matrix

$$= v^T S v \quad \dots \text{called } \mathbf{quadric} \text{ (also in N-D)}$$



PCA: data in **eigen coordinates** → variances are the **eigenvalues**



mean

$$\mu = E(x) = \int x p(x) dx$$

median m

$$\int_a^m p(x) dx = \frac{1}{2}$$

variance

$$\sigma^2 = \text{var}(x) = \int (x - \mu)^2 p(x) dx$$

$$\sigma^2 = E(x^2) - E(x)^2$$

$$\sigma_{tot}^2 = \sigma_1^2 + \sigma_2^2 + 2 \text{cov}(x_1, x_1)$$

covariance

$$\text{cov}(x_1, x_1) = E(x_1 x_2) - 2E(x_1)E(x_2)$$

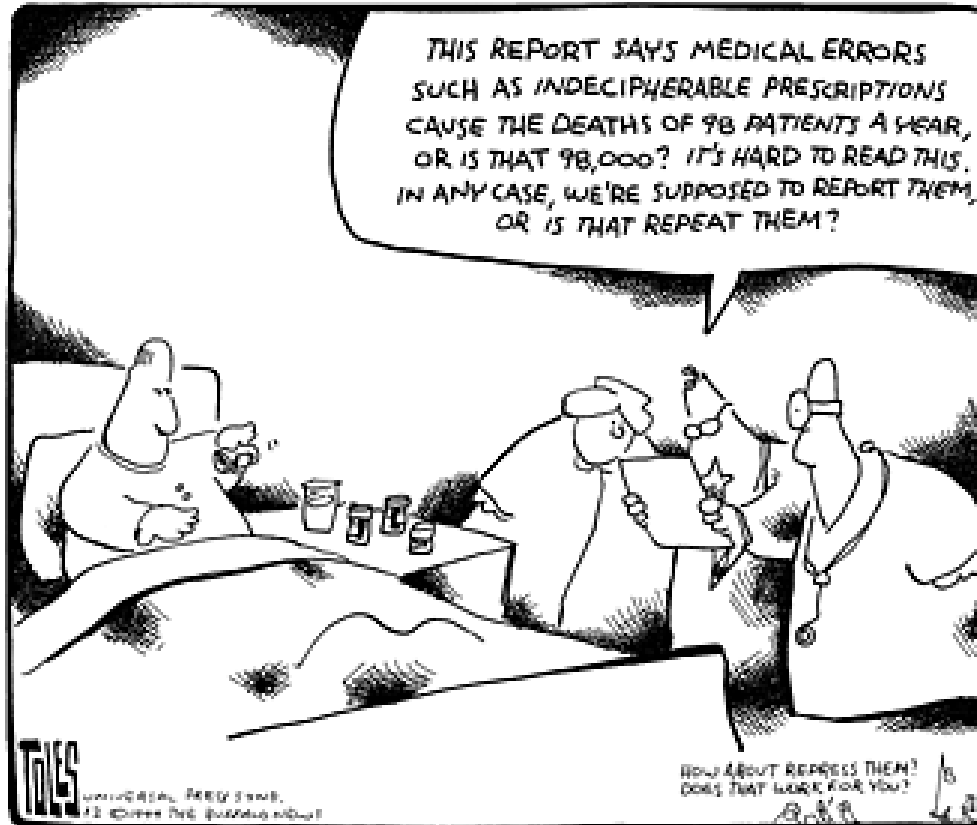
**correlation
coefficient**

$$\rho(x_1, x_1) = \frac{\text{cov}(x_1, x_1)}{\sqrt{\sigma_1^2 \sigma_2^2}}$$

note:

$$\int (x - \mu)^n p(x) dx$$

called nth *moment*
of a pdf



TOLES©1999 The Buffalo News. Reprinted with permission of UNIVERSAL PRESS SYNDICATE. All rights reserved.

Outline:

Basics

Most Common PDFs

- uniform
- binomial
- Poissonian
- Normal/Gaussian

Error Estimation

Bayesian Statistics



finding those $p(x)$ that maximize the **entropy S**, given constraints **C**

$$S = - \int p(x) \ln(x) dx \quad \text{c:} \quad \int p(x) dx = 1 \quad \longrightarrow \quad p(x) = \text{const}$$

$$\mu = \int_a^b x p(x) dx = \text{const} \int_a^b x dx = \text{const} \frac{1}{2} (b^2 - a^2)$$

$$\int_a^b p(x) dx = 1 \quad \text{const} \int_a^b dx = 1 \quad \rightarrow \text{const} = \frac{1}{b - a}$$

$$\mu = \frac{1}{2} \frac{b^2 - a^2}{b - a}$$

$$\sigma^2 = \frac{1}{12} (b - a)^2$$

Note: the uniform distribution has the largest entropy

→ maximum ignorance = no prior information = unbiased



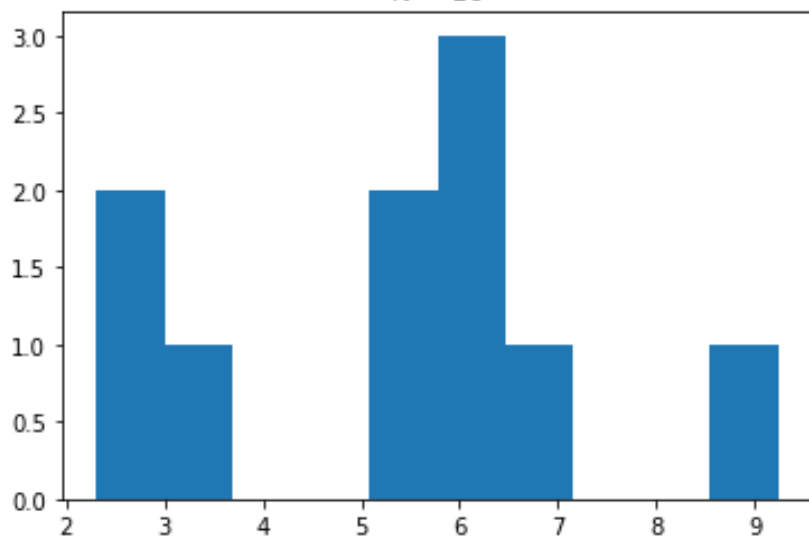
$$p(x) = \text{const}$$

plotting the pdf

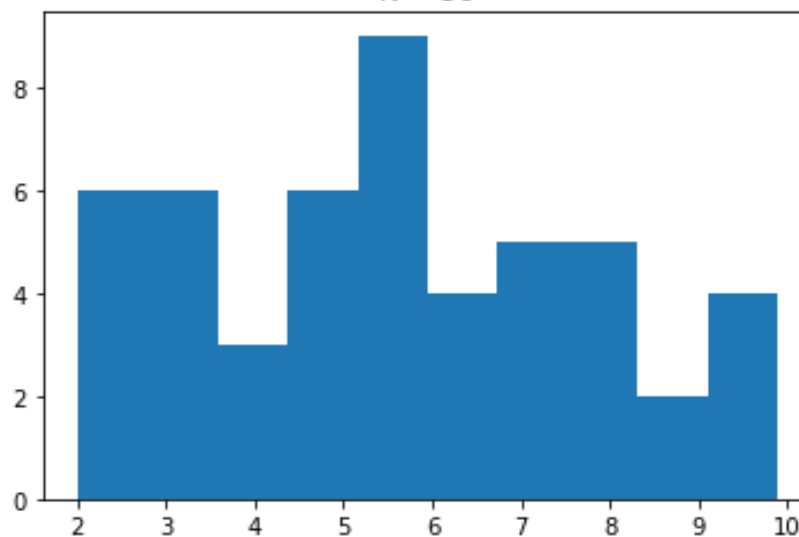
```
U = np.random.uniform(low, high, shape)  
plt.hist(U)
```

continuous support

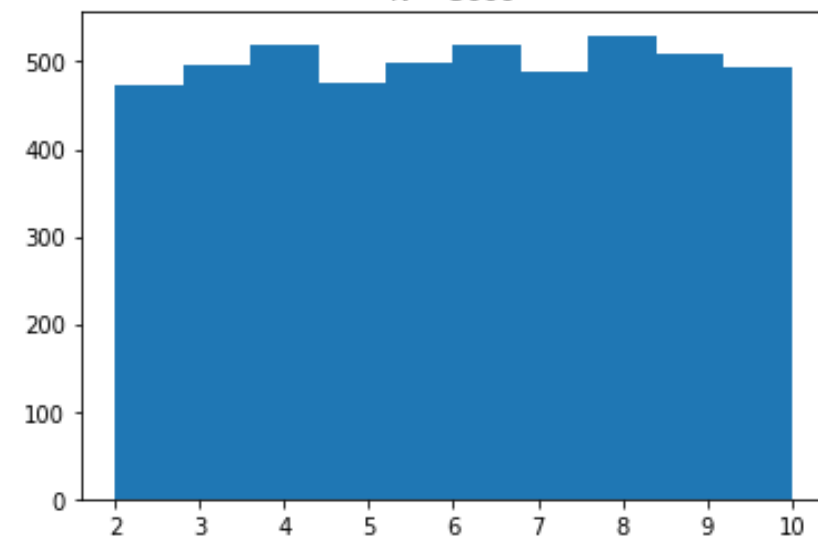
N = 10



N = 50



N = 5000





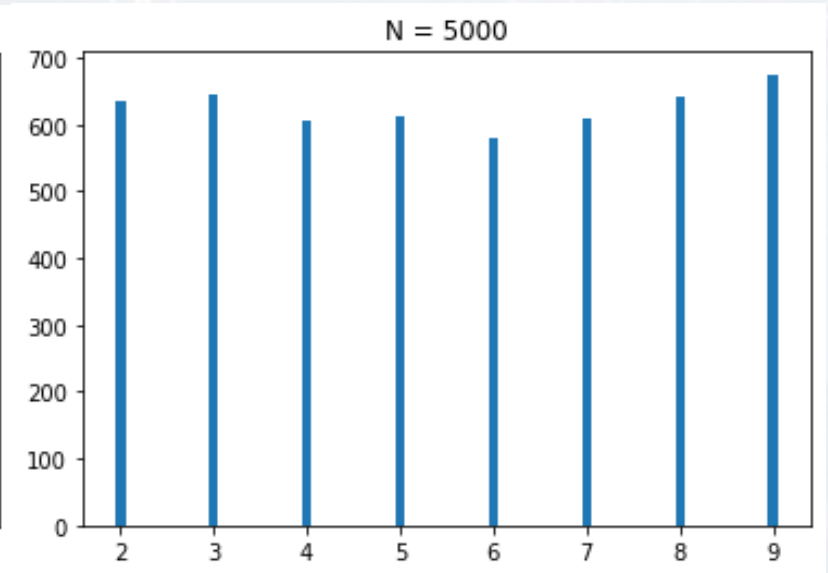
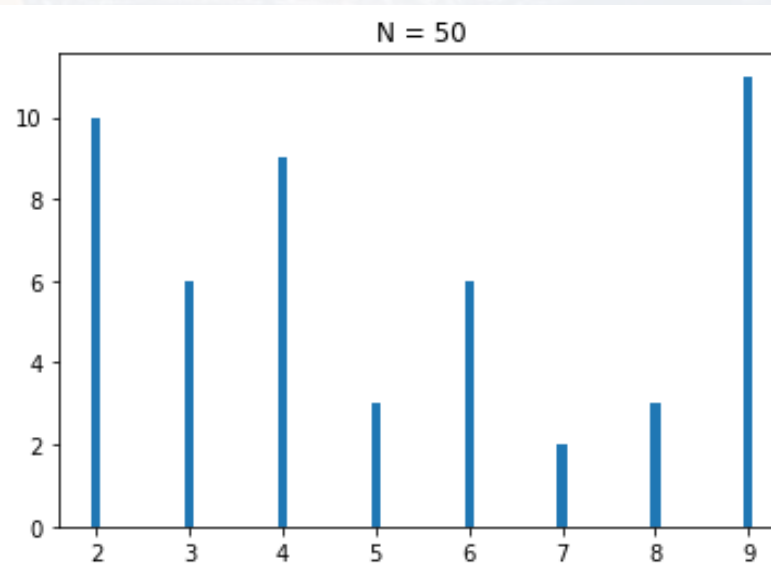
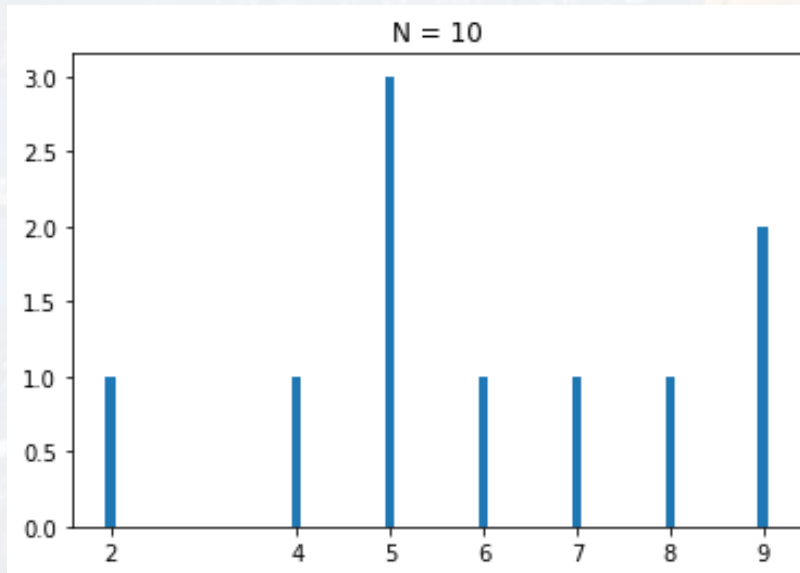
$$p(x) = \text{const}$$

plotting the pdf

```
U = np.random.randint(low, high, shape)
```

discrete support

```
labels, counts = np.unique(U, return_counts = True)
plt.bar(labels, counts, align = 'center', width = 0.1)
plt.gca().set_xticks(labels)
plt.title('N = ' + str(N))
```





$$p(x) = \text{const}$$

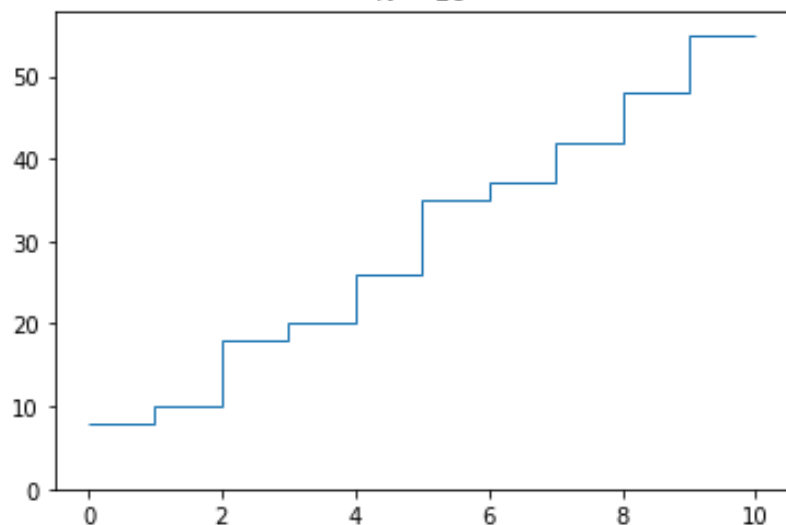
plotting the cdf

```
U = np.random.randint(low, high, shape)
```

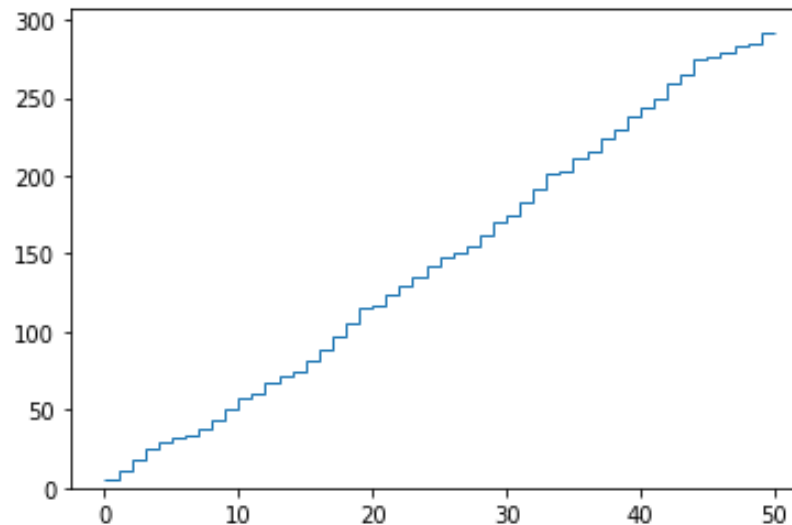
discrete support

```
C = np.cumsum(U)  
plt.stairs(C, baseline = None)  
plt.title('N = ' + str(N))
```

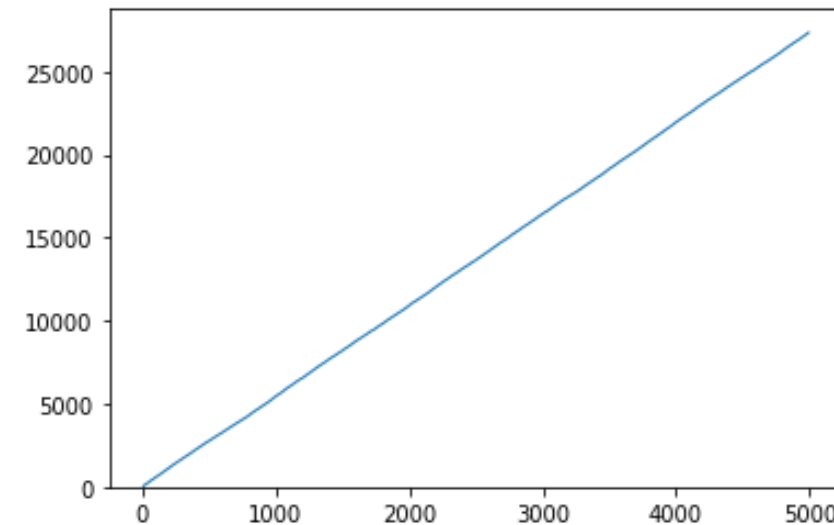
N = 10

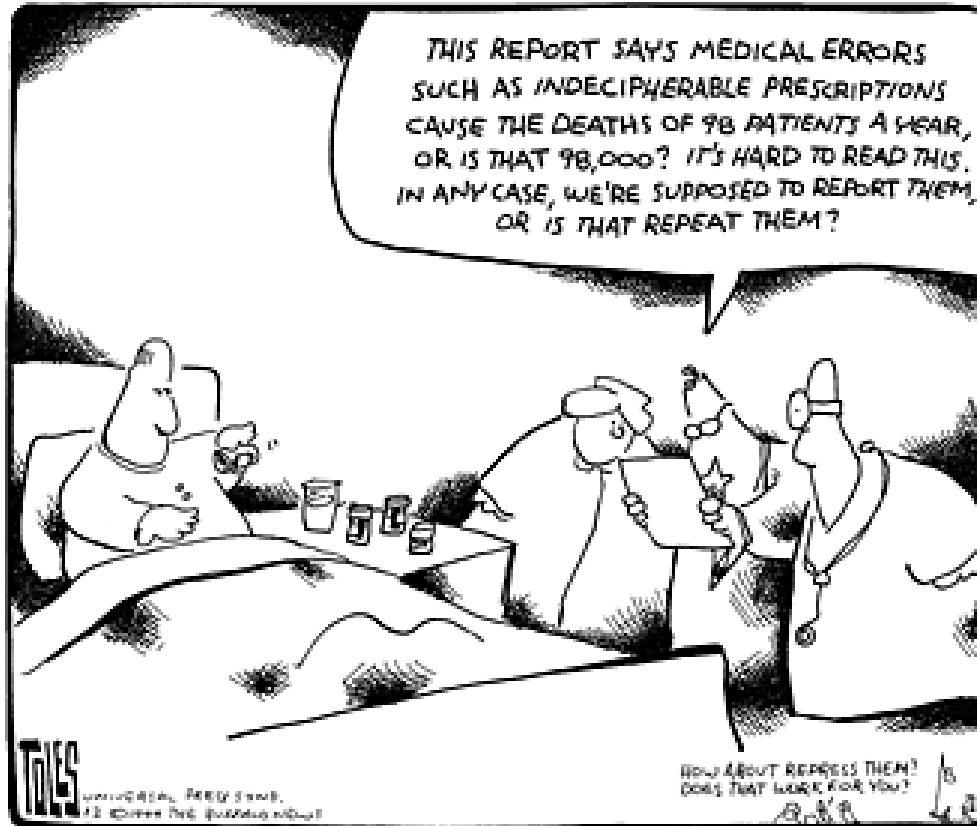


N = 50



N = 5000





TOLES©1999 The Buffalo News. Reprinted with permission of UNIVERSAL PRESS SYNDICATE. All rights reserved.

Outline:

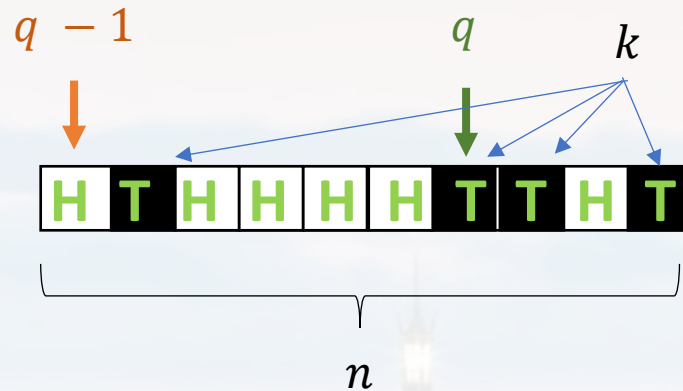
Basics

Most Common PDFs

- uniform
- **binomial**
- Poissonian
- Normal/Gaussian

Error Estimation

Bayesian Statistics



probability of having a sequence of **k tails** and **n-k heads**

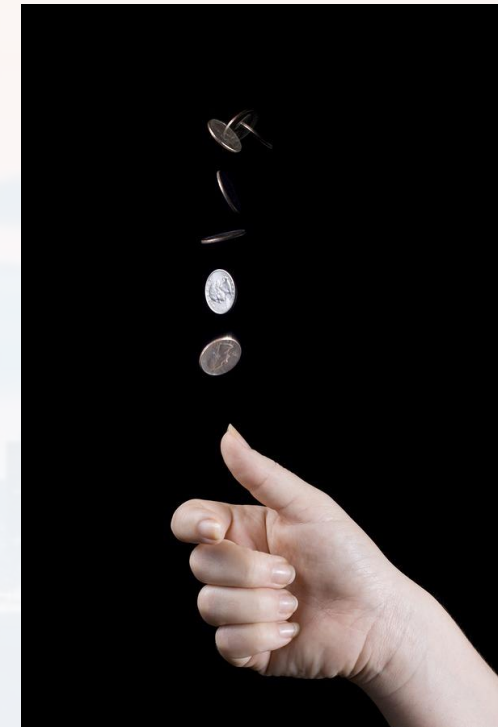
$$p_{tot} = \prod_i q_i^{n_i} = q^k (1 - q)^{n-k}$$

probability of having **any** sequence of **k tails** and **n-k heads**

$$P(k|q, n) = \binom{n}{k} q^k (1 - q)^{n-k}$$

binomial distribution

$$\frac{n!}{k!(n-k)!} =: \binom{n}{k} \quad \text{“n choose k”}$$



fair coin? $q = 0.5$???

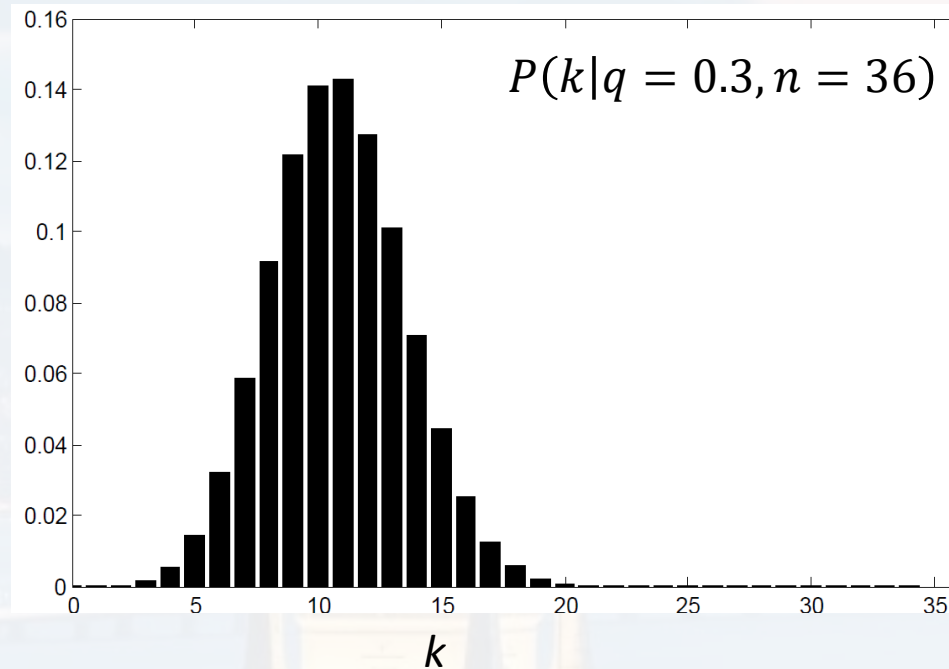


$$P(k|q, n) = \binom{n}{k} q^k (1 - q)^{n-k}$$

binomial distribution

$$\mu = \sum_i x_i p(x_i)$$

$$\mu = \int x p(x) dx$$



$$\mu = \sum_{k=0}^n k \binom{n}{k} q^k (1 - q)^{n-k} = qn$$

$$\text{var}(k) = \sum_{k=0}^n (k - qn)^2 \binom{n}{k} q^k (1 - q)^{n-k} = qn(1 - q)$$



$$P(k|q, n) = \binom{n}{k} q^k (1 - q)^{n-k}$$

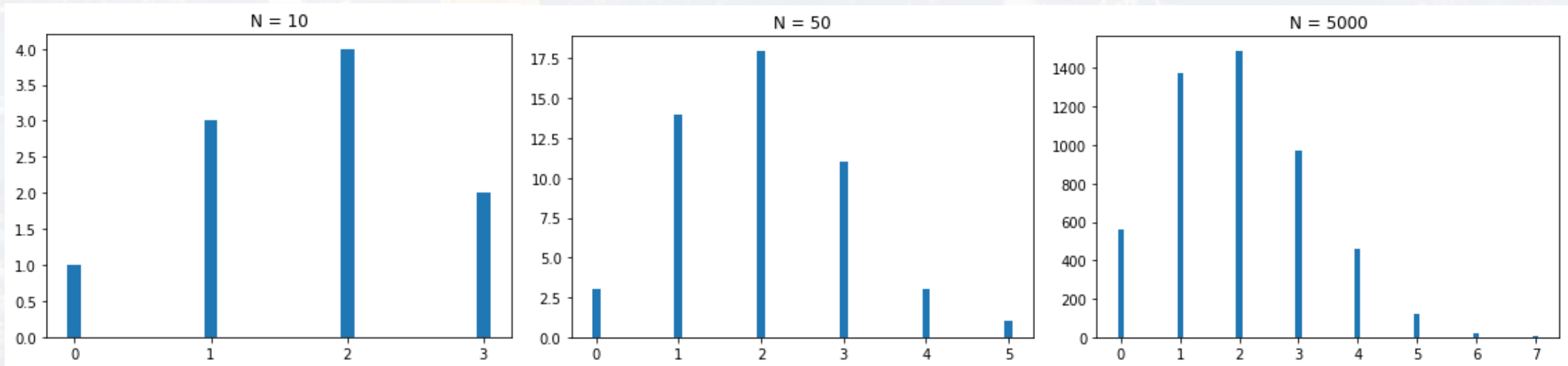
binomial distribution

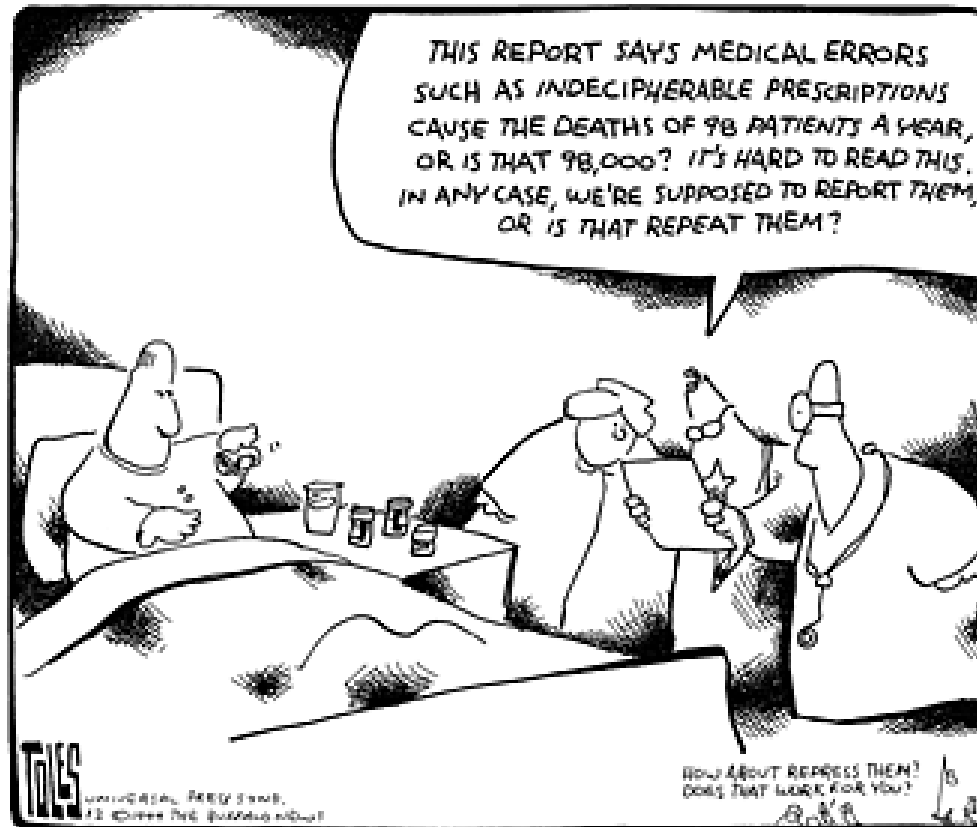
$q = 0.2$

$n = 10$

$K = \text{np.random.binomial}(n, q, N)$

```
labels, counts = np.unique(K, return_counts = True)
plt.bar(labels, counts, align = 'center', width = 0.1)
plt.gca().set_xticks(labels)
plt.title('N = ' + str(N))
```





TOLES©1999 The Buffalo News. Reprinted with permission of UNIVERSAL PRESS SYNDICATE. All rights reserved.

Outline:

Basics

Most Common PDFs

- uniform
- binomial
- **Poissonian**
- Normal/Gaussian

Error Estimation

Bayesian Statistics



$$P(k|q, n) = \binom{n}{k} q^k (1 - q)^{n-k}$$

binomial distribution

rare events

$$\rightarrow q \ll 1$$

Taylor expansion for $(1 - q)^{n-k}$ around $q = 0$

$$(1 - q)^{n-k} = 1 - nq + \frac{(nq)^2}{2} - \frac{(nq)^3}{6} + \dots = e^{-nq}$$

$$\rightarrow n \rightarrow \infty$$

Stirling's approximation for $n!$

$$\frac{n!}{(n-k)!} \approx \sqrt{\frac{n}{n-k}} \frac{n^n e^{n-k}}{e^n (n-k)^{n-k}} \approx n^k$$

$$\binom{n}{k} q^k (1 - q)^{n-k} \approx \frac{(nq)^k e^{-nq}}{k!}$$

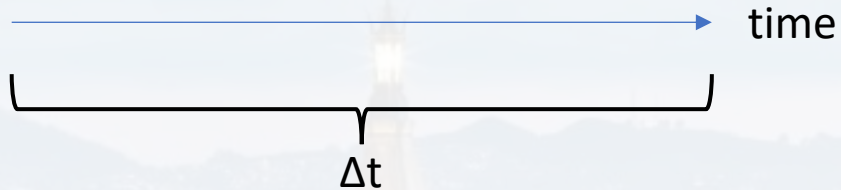


$$\binom{n}{k} q^k (1 - q)^{n-k} \approx \frac{(nq)^k e^{-nq}}{k!}$$

often: $nq := \lambda$

events per time interval: $\lambda = c \Delta t$

H T H H H H T T H T



rate $c = 4$ tails per Δt

$$P(k|c) = \frac{(c \Delta t)^k e^{-c \Delta t}}{k!}$$

Poisson distribution

$$\mu = qn \rightarrow qn = \lambda$$

$$\text{var}(k) = qn(1 - q) \rightarrow qn = \lambda$$





$$P(k|c) = \frac{(c \Delta t)^k e^{-c \Delta t}}{k!}$$

Poisson distribution

$$\mu = \lambda$$

$$\text{var}(k) = \lambda$$

- **rare** events
- events are mutually **independent**
- events have **no duration**

examples:

- radioactive decay
- single photon detection
- lightning
- mutation of a gene
- receiving WhatsApp messages/SMS



$$P(k|c) = \frac{(c \Delta t)^k e^{-c \Delta t}}{k!}$$

Poisson distribution

$$\mu = \lambda$$

$$\text{var}(k) = \lambda$$

```
c      = 5  
delt   = 10  
lam    = c * delt
```

```
K      = np.random.poisson(lam, N)
```

```
labels, counts = np.unique(K, return_counts = True)  
plt.bar(labels, counts, align = 'center', width = 0.1)  
plt.gca().set_xticks(labels)  
plt.title('N = ' + str(N))
```



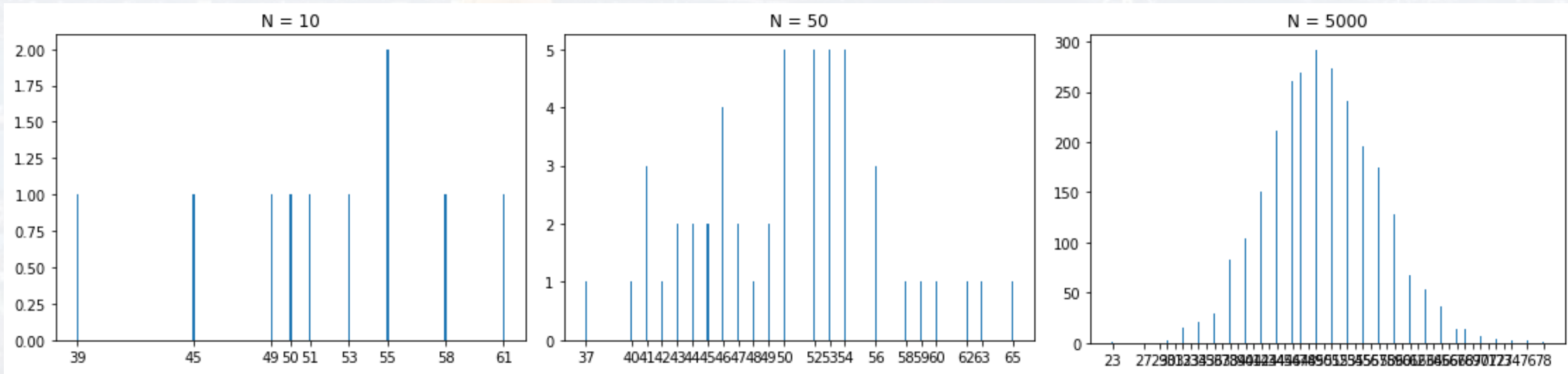

$$P(k|c) = \frac{(c \Delta t)^k e^{-c \Delta t}}{k!}$$

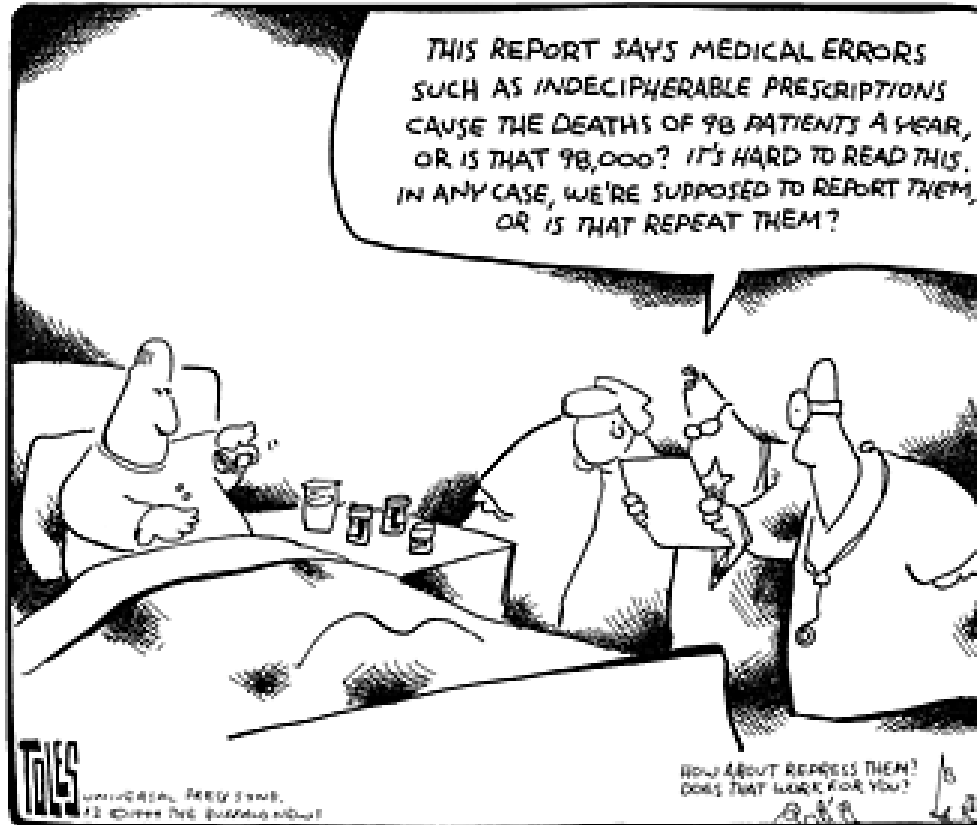
Poisson distribution

$$\mu = \lambda$$

$$\text{var}(k) = \lambda$$

```
c = 5
delt = 10
lam = c * delt
K = np.random.poisson(lam, N)
labels, counts = np.unique(K, return_counts = True)
plt.bar(labels, counts, align = 'center', width = 0.1)
plt.gca().set_xticks(labels)
plt.title('N = ' + str(N))
```





TOLES©1999 The Buffalo News. Reprinted with permission of UNIVERSAL PRESS SYNDICATE. All rights reserved.

Outline:

Basics

Most Common PDFs

- uniform
- binomial
- Poissonian
- **Normal/Gaussian**

Error Estimation

Bayesian Statistics

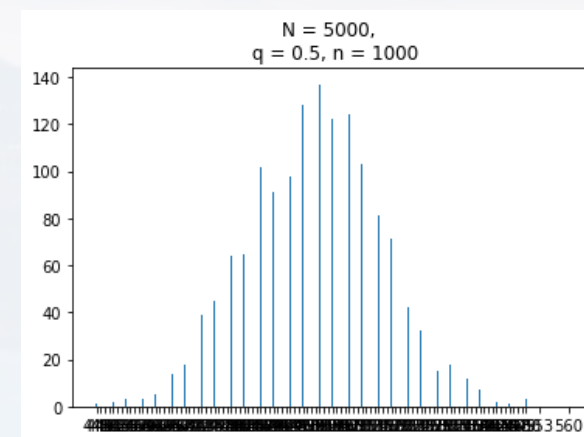
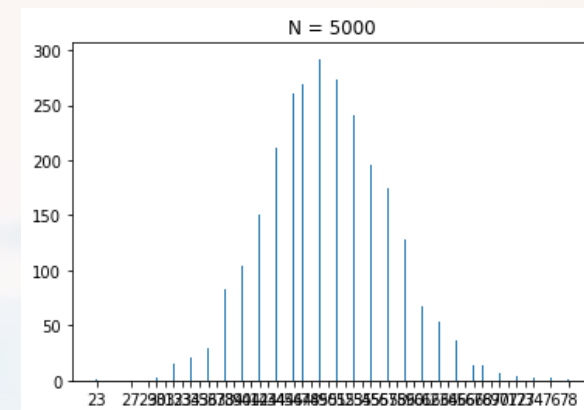


$$P(k|c) = \frac{(c \Delta t)^k e^{-c \Delta t}}{k!}$$

Poisson distribution

$$P(k|q, n) = \binom{n}{k} q^k (1 - q)^{n-k}$$

binomial distribution



Stirling's approximation for even larger n

$$P(k|n, p) \approx \frac{1}{\sqrt{2\pi nq(1-q)}} \exp\left[-\frac{(k - nq)^2}{2nq(1-q)}\right]$$



Stirling's approximation for even larger n

$$P(k|n, p) \approx \frac{1}{\sqrt{2\pi nq(1-q)}} \exp\left[-\frac{(k-nq)^2}{2nq(1-q)}\right]$$

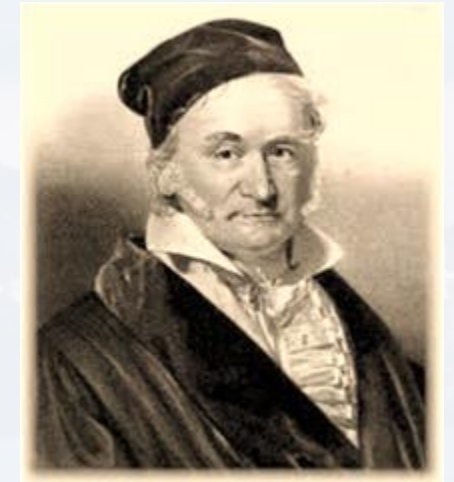
using $\sigma^2 = \text{var}(k) = qn(1-q)$

$$\mu = qn$$

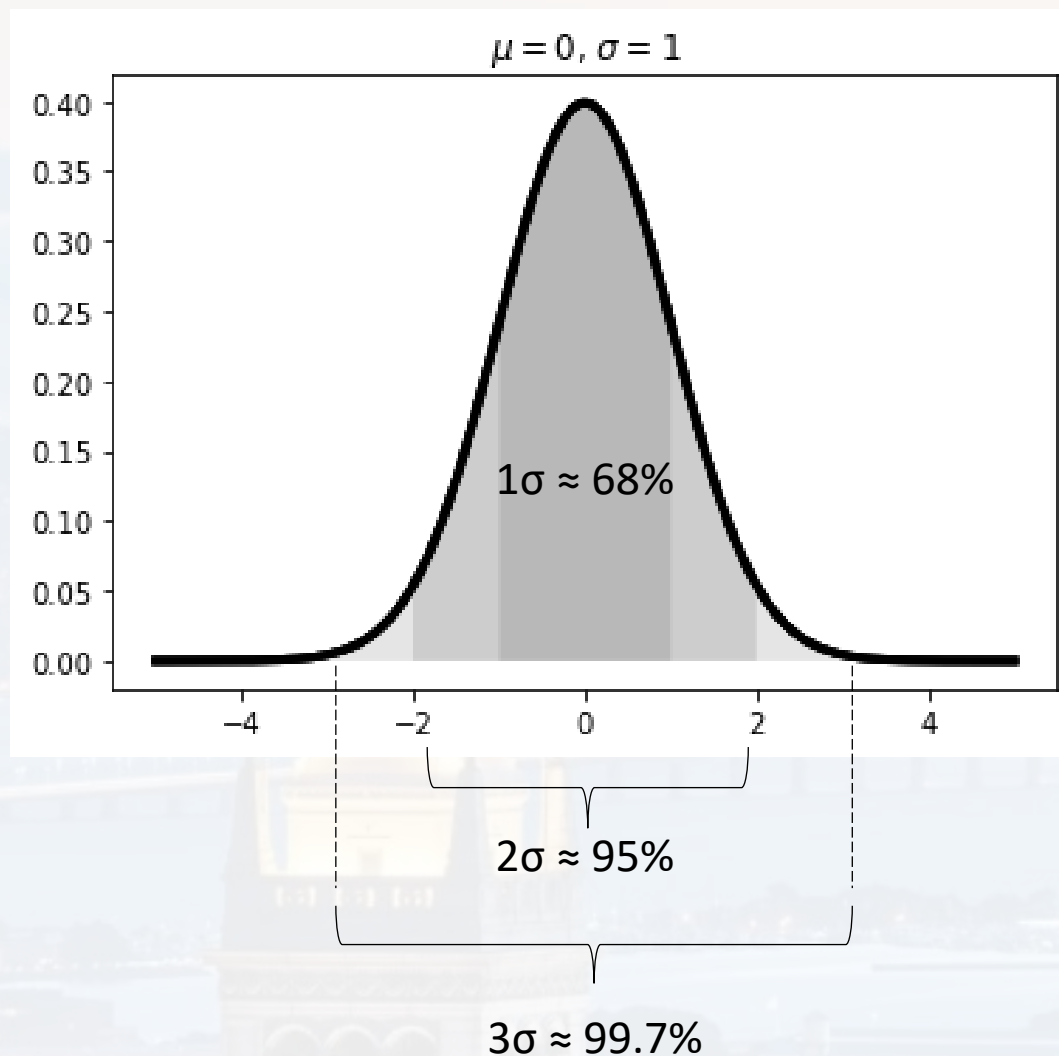
and $k := x$

$$P(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi \sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2 \sigma^2}\right]$$

Normal/Gauss distribution

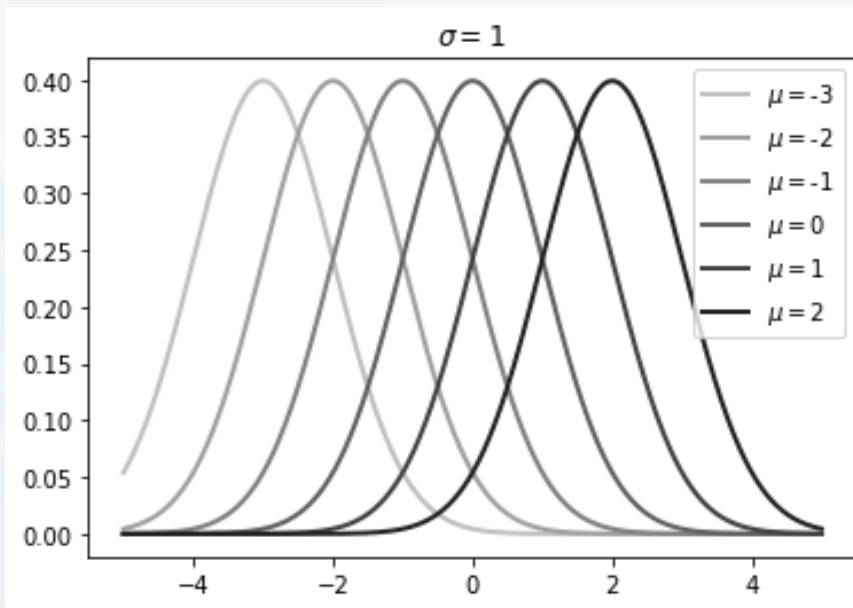


Note, that the **Poisson** and the **Binomial distribution** are *discrete*,
whereas the **Normal distribution** is *continuous*!



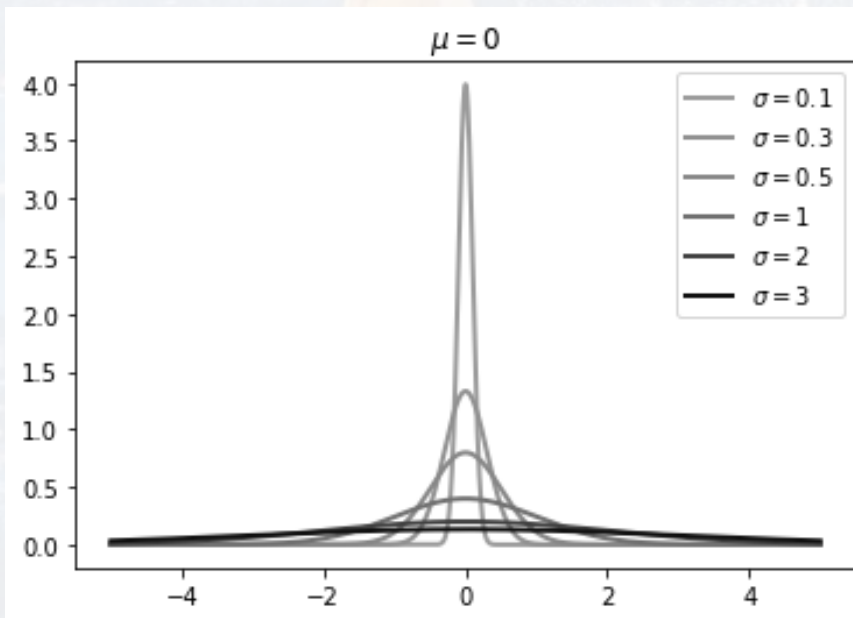
$$P(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi} \sigma} \exp \left[-\frac{(x - \mu)^2}{2 \sigma^2} \right]$$

Normal/Gauss distribution



$$P(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi} \sigma} \exp \left[-\frac{(x - \mu)^2}{2 \sigma^2} \right]$$

Normal/Gauss distribution

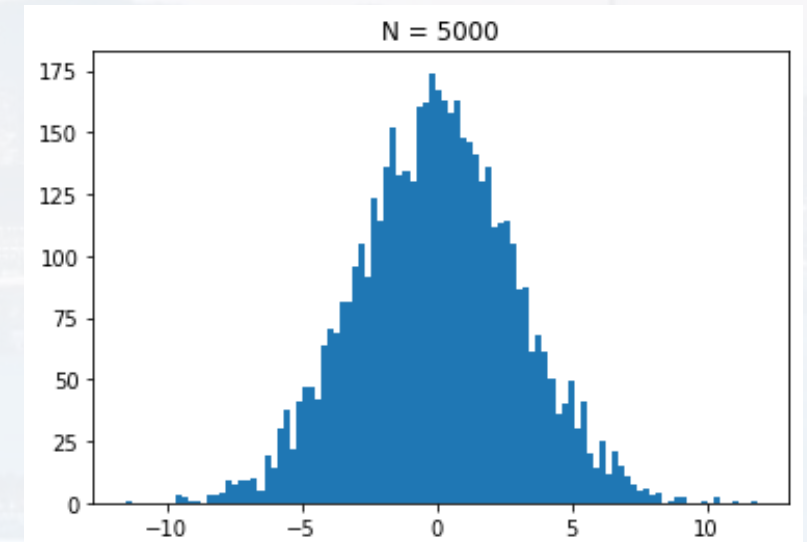
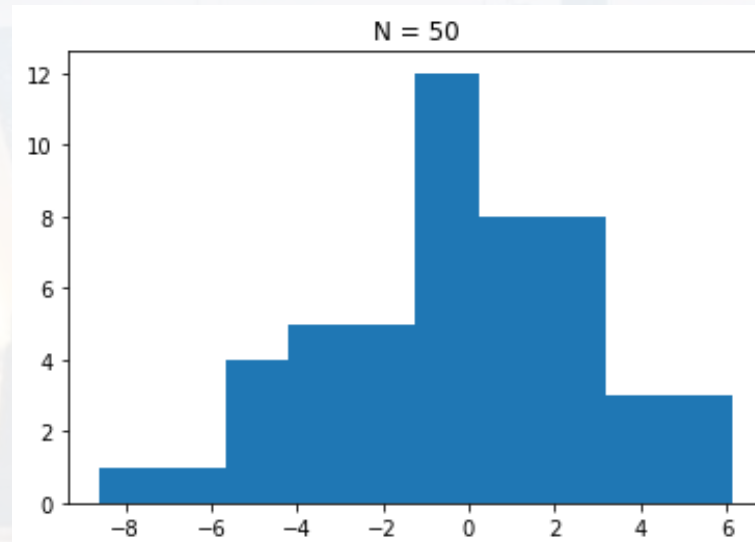
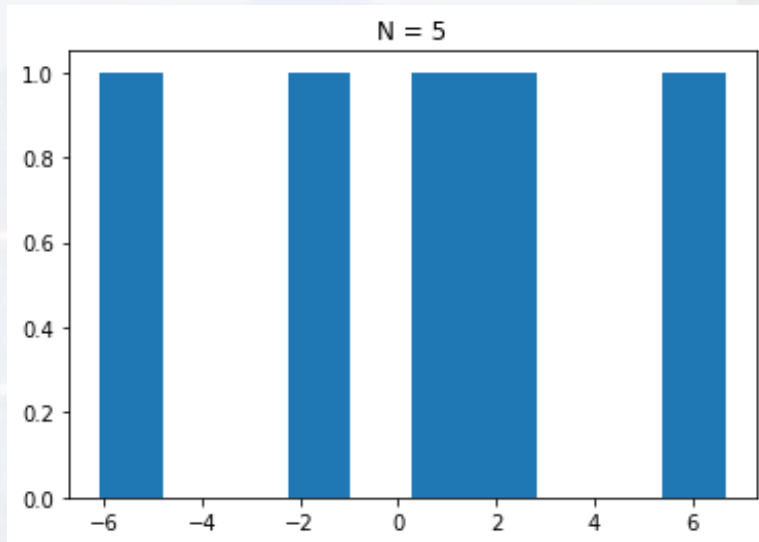


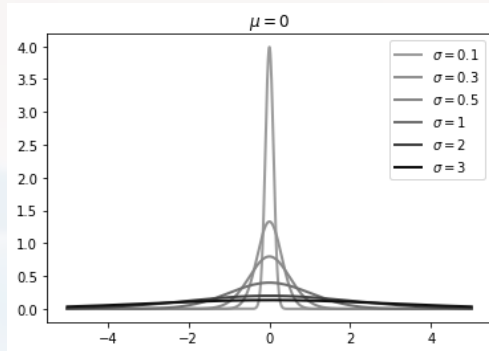


```
mu = 0  
s = 1  
P = np.random.normal(mu, s, N)  
plt.hist(P)  
plt.title('N = ' + str(N))
```

$$P(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi} \sigma} \exp \left[-\frac{(x - \mu)^2}{2 \sigma^2} \right]$$

Normal/Gauss distribution





$$P(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi} \sigma} \exp \left[-\frac{(x - \mu)^2}{2 \sigma^2} \right]$$

Normal/Gauss distribution

examples:

- diffusion processes
- approx. stat. error of data points
- approx. distribution of body height/shoe sizes/ weight, IQ
- approx. blood pressure, blood values
- approx. retirement age
-

applications:

- significance tests
- t-test
- ANOVA/MANOVA
- χ^2 - test
- χ^2 itself and students-t distribution
- ...

Why do so many quantities follow a normal distribution?



Why do so many quantities follow a normal distribution?

At the end... all probability distributions are **Maximum Entropy** Distributions, subject to a **set of constraints**

Distribution name	Probability density / mass function	Maximum Entropy constraint	Support
Uniform (discrete)	$f(k) = \frac{1}{b - a + 1}$	None	$\{a, a + 1, \dots, b - 1, b\}$
Uniform (continuous)	$f(x) = \frac{1}{b - a}$	None	$[a, b]$
Bernoulli	$f(k) = p^k (1 - p)^{1-k}$	$\mathbb{E}[K] = p$	$\{0, 1\}$
Geometric	$f(k) = (1 - p)^{k-1} p$	$\mathbb{E}[K] = \frac{1}{p}$	$\mathbb{N} \setminus \{0\} = \{1, 2, 3, \dots\}$
Exponential	$f(x) = \lambda \exp(-\lambda x)$	$\mathbb{E}[X] = \frac{1}{\lambda}$	$[0, \infty)$
Laplace	$f(x) = \frac{1}{2b} \exp\left(-\frac{ x - \mu }{b}\right)$	$\mathbb{E}[X - \mu] = b$	$(-\infty, \infty)$
Asymmetric Laplace	$f(x) = \frac{\lambda \exp\left(-(x - m) \lambda s \kappa^s\right)}{\left(\kappa + \frac{1}{\kappa}\right)}$ where $s \equiv \text{sgn}(x - m)$	$\mathbb{E}[(X - m) s \kappa^s] = \frac{1}{\lambda}$	$(-\infty, \infty)$
Pareto	$f(x) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}}$	$\mathbb{E}[\ln X] = \frac{1}{\alpha} + \ln(x_m)$	$[x_m, \infty)$
Normal	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$	$\mathbb{E}[X] = \mu,$ $\mathbb{E}[X^2] = \sigma^2 + \mu^2$	$(-\infty, \infty)$



Why do so many quantities follow a normal distribution?

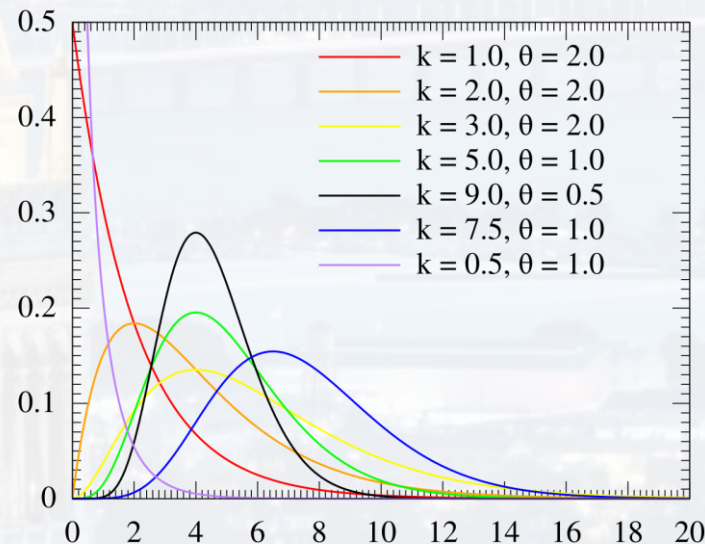
At the end... all probability distributions **are Maximum Entropy** Distributions, subject to a **set of constraints**

examples:

- approx. stat. error of data points
- approx. distribution of body height/shoe sizes/ weight, IQ
- approx. blood pressure, blood values
- approx. retirement age

....

Gamma	$f(x) = \frac{x^{k-1} \exp\left(-\frac{x}{\theta}\right)}{\theta^k \Gamma(k)}$	$\begin{aligned}\mathbb{E}[X] &= k\theta, \\ \mathbb{E}[\ln X] &= \psi(k) + \ln \theta\end{aligned}$	$[0, \infty)$
-------	--	--	---------------





binomial distribution

$$P(k|q, n) = \binom{n}{k} q^k (1 - q)^{n-k}$$

$q \rightarrow 0$

Poisson distribution

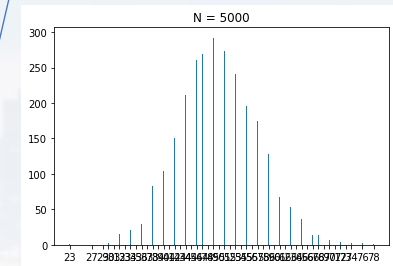
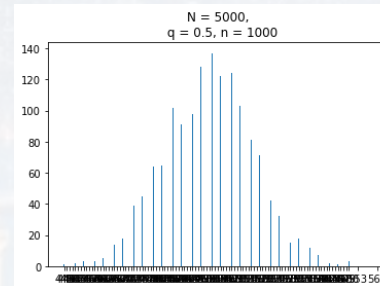
$$P(k|c) = \frac{(c \Delta t)^k e^{-c \Delta t}}{k!}$$

$n \rightarrow \infty$

$n \rightarrow \infty$

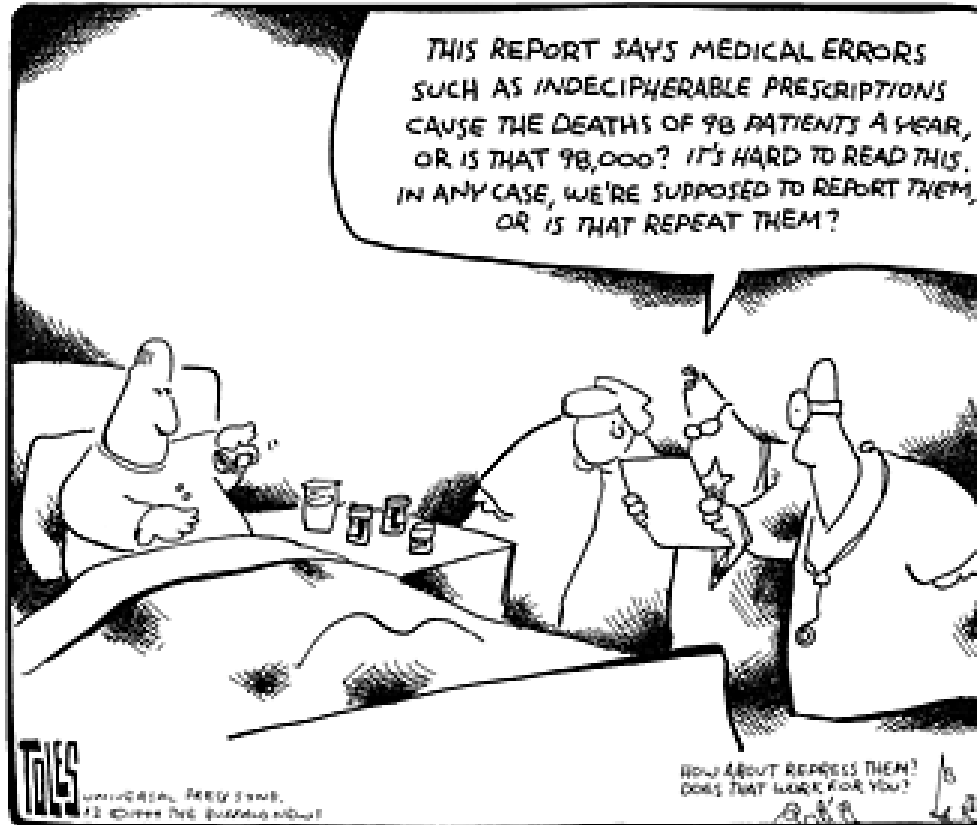
Normal/Gauss distribution

$$P(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi} \sigma} \exp \left[-\frac{(x - \mu)^2}{2 \sigma^2} \right]$$



The fact that that many datasets can be well approximated by a Normal distribution for $n \rightarrow \infty$ is called

Central Limit Theorem



TOLES©1999 The Buffalo News. Reprinted with permission of UNIVERSAL PRESS SYNDICATE. All rights reserved.

Outline:

Basics

Most Common PDFs

- uniform
- binomial
- Poissonian
- Normal/Gaussian

Error Estimation

Bayesian Statistics