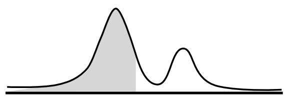
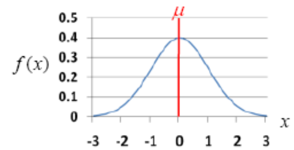
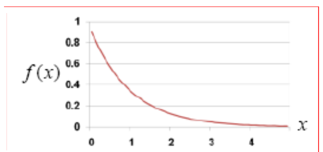
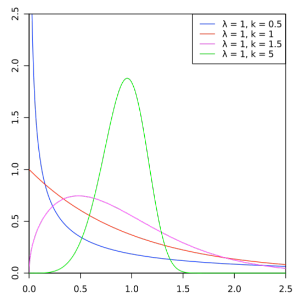




continuous random variables





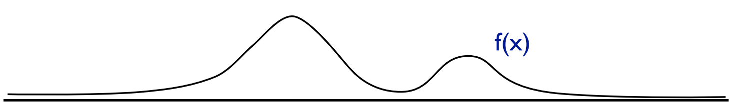
Discrete random variable: takes values in a finite or countable set, e.g.

- $X \in \{1, 2, \ldots, 6\}$ with equal probability
- X is positive integer i with probability 2^{-i}

Continuous random variable: takes values in an uncountable set, e.g.

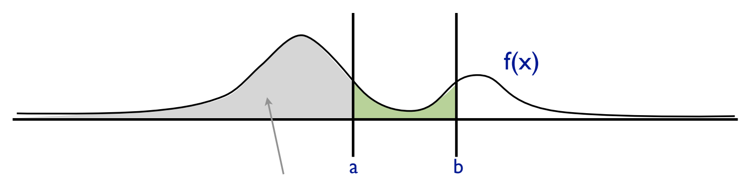
- X is the weight of a random person (a real number)
- X is a randomly selected point inside a unit square
- X is the waiting time until the next packet arrives at the server

$f(x): \mathbb{R} \to \mathbb{R}$, the *probability density function* (or simply "density")



Require:	I.e., distribution is:
$f(x) \geq 0$, and	← nonnegative, and
$\int_{-\infty}^{+\infty} f(x)\,dx = 1$	← normalized,
	just like discrete PMF

F(x): the *cumulative distribution function* (aka the "distribution")

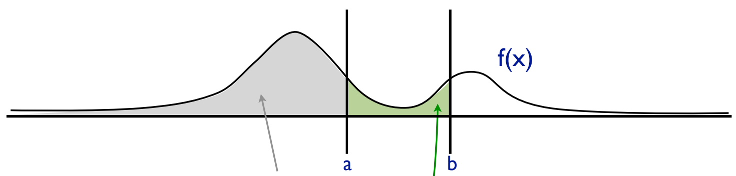


$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)\,dx$ (Area left of a)

$P(a < X \leq b) =$

cdf

F(x): the *cumulative distribution function* (aka the "distribution")

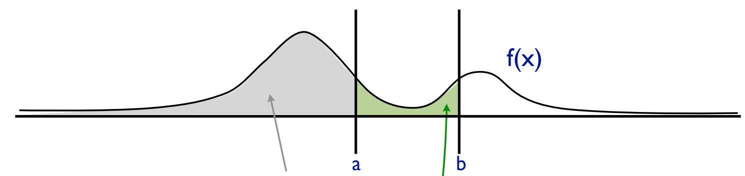


$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)\,dx$ (Area left of a)

$P(a < X \leq b) = F(b) - F(a)$ (Area between a and b)

cdf

F(x): the *cumulative distribution function* (aka the "distribution")



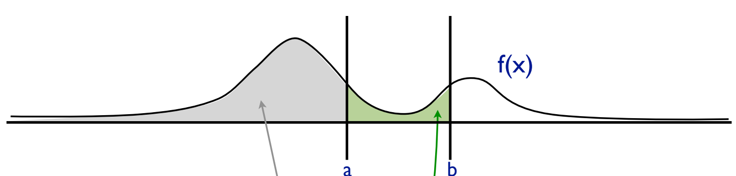
$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)\,dx$ (Area left of a)

$P(a < X \leq b) = F(b) - F(a)$ (Area between a and b)

Relationship between f(x) and F(x)?

cdf

F(x): the *cumulative distribution function* (aka the "distribution")



$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)\,dx$ (Area left of a)

$P(a < X \leq b) = F(b) - F(a)$ (Area between a and b)

A key relationship:

$f(x) = \frac{d}{dx} F(x)$, since $F(a) = \int_{-\infty}^{a} f(x)\,dx$,

why is it called a density?

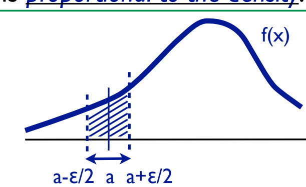
Densities are *not* probabilities; e.g. may be > 1

$P(X = a) = \lim_{\varepsilon \to 0} P(a-\varepsilon < X \leq a) = F(a)-F(a) = 0$

I.e., the probability that a continuous r.v. falls *at* a specified point is *zero*.

But

the probability that it falls *near* that point is *proportional to the density*:



why is it called a density?

Densities are *not* probabilities; e.g. may be > 1

$P(X = a) = \lim_{\varepsilon \to 0} P(a-\varepsilon < X \leq a) = F(a)-F(a) = 0$

I.e.,

the probability that a continuous r.v. falls *at* a specified point is *zero*.

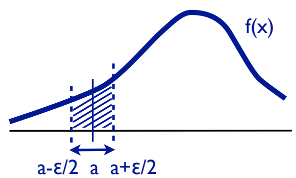
But

the probability that it falls *near* that point is *proportional to the density*:

$P(a - \varepsilon/2 < X \leq a + \varepsilon/2) =$

$F(a + \varepsilon/2) - F(a - \varepsilon/2)$

$\approx \varepsilon \cdot f(a)$



I.e., in a large random sample, expect more samples where density is higher (hence the name "density").

sums and integrals; expectation

Much of what we did with discrete r.v.s carries over almost unchanged, with Σ_x... replaced by $\int$... dx

E.g.

For discrete r.v. X, $E[X] = \Sigma_x \, xp(x)$

For continuous r.v. X, $E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$

sums and integrals; expectation

Much of what we did with discrete r.v.s carries over almost unchanged, with Σ_x... replaced by $\int$... dx

E.g.

For discrete r.v. X, $E[X] = \Sigma_x \, xp(x)$

For continuous r.v. X, $E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$

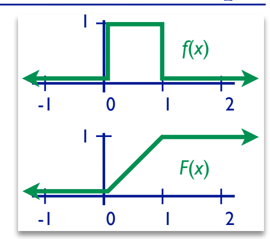
Why?

(a) We define it that way

(b) The probability that X falls "near" x, say within $x \pm dx/2$, is $\approx f(x)dx$, so the "average" X should be $\approx \Sigma\, xf(x)dx$ (summed over grid points spaced dx apart on the real line) and the limit of that as $dx \to 0$ is $\int xf(x)dx$

example

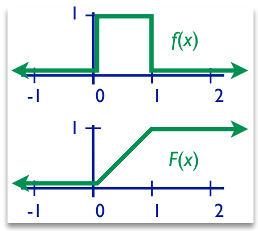
Let $f(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$



What is F(x)? What is E(X)?

example

Let $f(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$



$$\begin{aligned}
F(a) &= \int_{-\infty}^{a} f(x)dx \\
&= \begin{cases} 0 & \text{if } a \le 0 \\ a & \text{if } 0 < a \le 1 \ (\text{since } a = \int_0^a 1dx) \\ 1 & \text{if } 1 < a \end{cases} \\
E[X] &= \int_{-\infty}^{\infty} xf(x)dx = \int_0^1 x\,dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2} \\
E[X^2] &= \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^1 x^2\,dx = \left.\frac{x^3}{3}\right|_0^1 = \frac{1}{3}
\end{aligned}$$

properties of expectation

Linearity

E[aX+b] = aE[X]+b

E[X+Y] = E[X]+E[Y]

still true, just as for discrete

Functions of a random variable

E[g(X)] = ∫g(x)f(x)dx

just as for discrete, but w/integral

Alternatively, let Y = g(X), find the density of Y, say f_Y, and directly compute E[Y] = ∫$yf_Y(y)$dy.

variance

Definition is same as in the discrete case

Var[X] = E[(X-μ)^2] where μ = E[X]

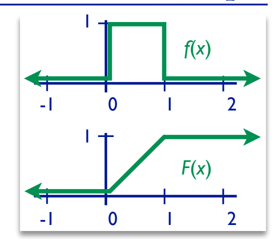
Identity still holds:

Var[X] = E[X^2] - (E[X])^2

proof "same"

example

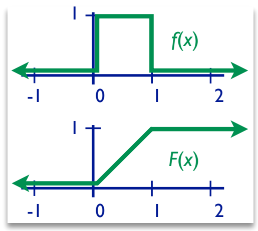
Let $f(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$



$$\begin{aligned}
F(a) &= \int_{-\infty}^{a} f(x)dx \\
&= \begin{cases} 0 & \text{if } a \le 0 \\ a & \text{if } 0 < a \le 1 \ (\text{since } a = \int_0^a 1dx) \\ 1 & \text{if } 1 < a \end{cases} \\
E[X] &= \int_{-\infty}^{\infty} xf(x)dx = \int_0^1 x\,dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2} \\
E[X^2] &= \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^1 x^2\,dx = \left.\frac{x^3}{3}\right|_0^1 = \frac{1}{3}
\end{aligned}$$

example

Let $f(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$



$$F(a) = \int_{-\infty}^{a} f(x)dx$$

$$= \begin{cases} 0 & \text{if } a \le 0 \\ a & \text{if } 0 < a \le 1 \text{ (since } a = \int_0^a 1dx) \\ 1 & \text{if } 1 < a \end{cases}$$

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_0^1 x\,dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2}$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^1 x^2\,dx = \left.\frac{x^3}{3}\right|_0^1 = \frac{1}{3}$$

$$\text{Var}[X] = E[X^2] - (E[X])^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \quad (\sigma \approx 0.29)$$

continuous random variables: summary

Continuous random variable X has density f(x), and

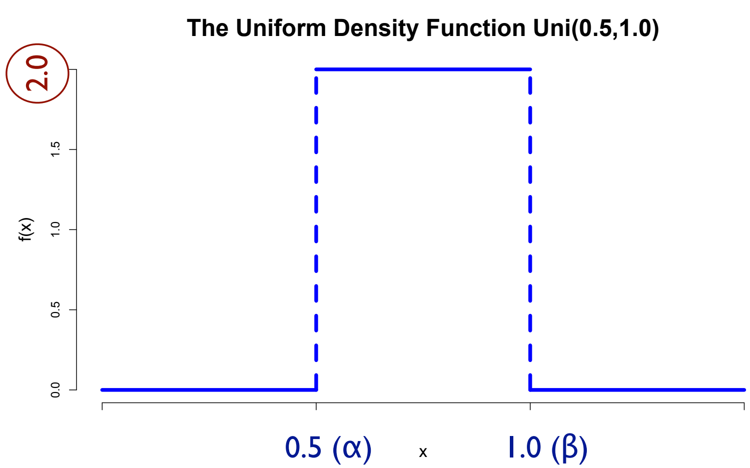
$$\Pr(a \le X \le b) = \int_a^b f(x)\,dx$$

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx$$

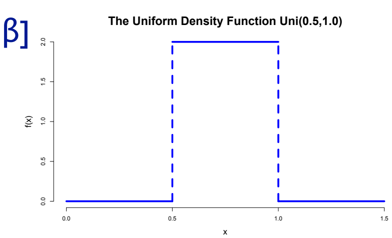
uniform random variables

X ~ Uni(α,β) is uniform in [α,β] $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha,\beta] \\ 0 & \text{otherwise} \end{cases}$



uniform random variables

X ~ Uni(α,β) is uniform in [α,β]



$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha,\beta] \\ 0 & \text{otherwise} \end{cases}$$

$$\Pr(a \le X \le b) = \int_a^b f(x)\,dx = \frac{b-a}{\beta-\alpha}$$

if α ≤ a ≤ b ≤ β:

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx = \frac{\alpha+\beta}{2}$$

Yes, you should review your basic calculus; e.g., these 2 integrals would be good practice.

Radioactive decay: How long until the next alpha particle?

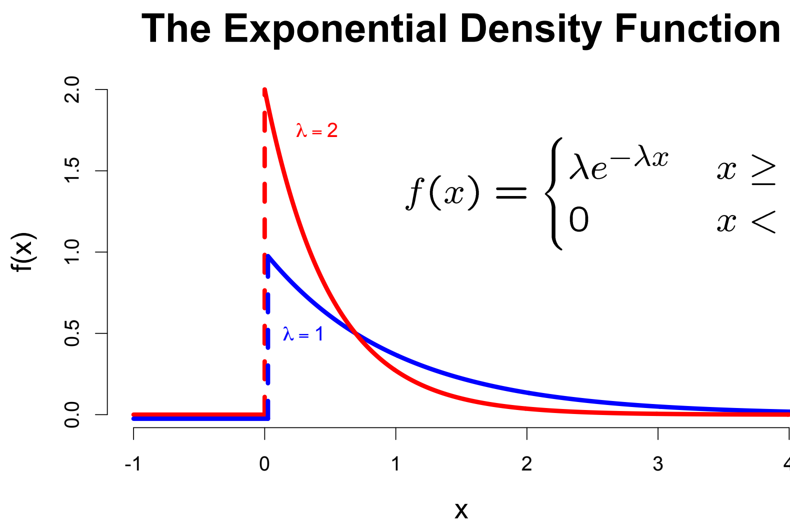
Customers: how long until the next customer/packet arrives at the checkout stand/server?

Buses: How long until the next #71 bus arrives on the Ave?

Yes, they have a schedule, but given the vagaries of traffic, riders with-bikes-and-baby-carriages, etc., can they stick to it?

Assuming events are independent, happening at some fixed *average* rate of λ per unit time – the waiting time until the next event is exponentially distributed (next slide)

X ~ Exp(λ)



$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

X ~ Exp(λ)

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

$$E[X] = \frac{1}{\lambda} \qquad \mathrm{Var}[X] = \frac{1}{\lambda^2}$$

$$\Pr(X \ge t) = e^{-\lambda t} = 1\text{-}F(t)$$

Memorylessness:

$$\Pr(X > s + t \mid X > s) = \Pr(X > t)$$

Assuming exp distr, if you've waited *s* minutes, prob of waiting *t* more is exactly same as *s = 0*

Same process, different measures:

Poisson: *how many* events in a *fixed time*;
Exponential: *how long* until the *next event*

λ is avg # per unit time;
1/λ is mean wait

Time it takes to check someone out at a grocery store is exponential with with an average value of 10.
If someone arrives to the line immediately before you, what is the probability that you will have to wait between 10 and 20 minutes?

$$T \sim exp(10^{-1})$$

$$Pr(10 \leq T \leq 20) = \int_{10}^{20} \frac{1}{10} e^{-\frac{x}{10}} dx$$

$$y = \frac{x}{10} \qquad dy = \frac{1}{10} dx$$

$$Pr(10 \leq T \leq 20) = \int_{1}^{2} e^{-y} dy = -e^{-y}\Big|_1^2 = (e^{-1} - e^{-2})$$

Suppose that the number of miles that a car can run before its battery wears out is exponentially distributed with an average value of 10,000 miles. If the car has already been used for 2000 miles, and the owner wants to take a 5000 mile trip, what is the probability she will be able to complete the trip without replacing the battery?

$$N \sim exp(1/10,000)$$

$$Pr(N \geq 7000 | N \geq 2000) = \frac{Pr(N \geq 7000)}{Pr(N \geq 2000)}$$

$$Pr(N \geq 7000) = e^{-7000/10000}$$

$$Pr(N \geq 2000) = e^{-2000/10000}$$

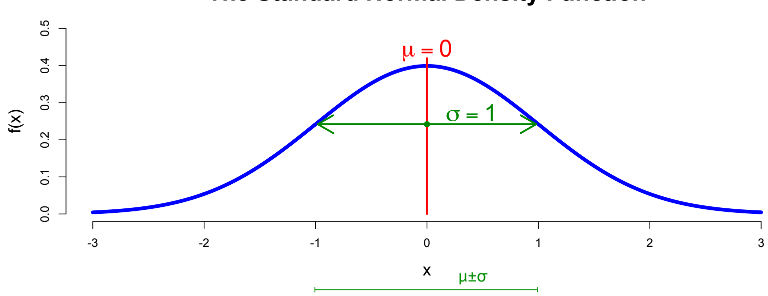
$$\text{answer} = e^{-5000/10000} = e^{-0.5}$$

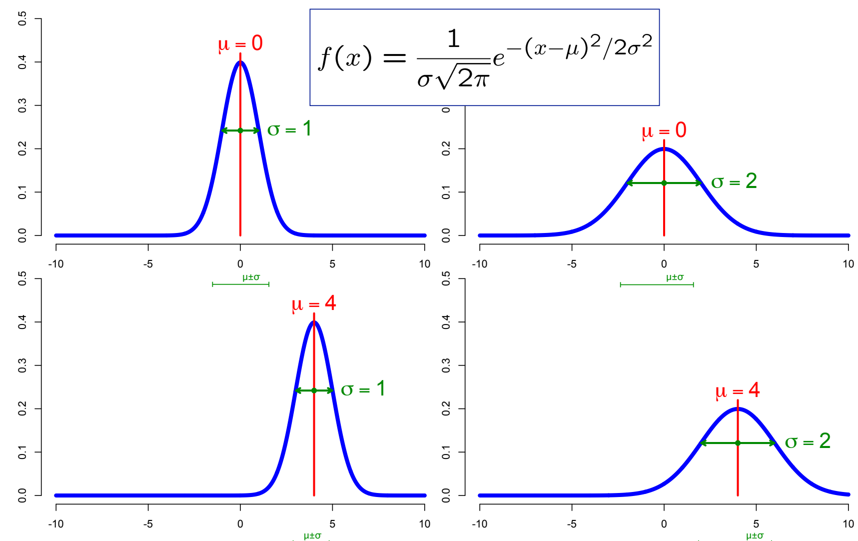
X is a normal (aka Gaussian) random variable X ~ N(μ, σ²)

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

$$E[X] = \mu \qquad \text{Var}[X] = \sigma^2$$

The Standard Normal Density Function





density at μ is ≈ .399/σ

normal random variables

X is a normal random variable   X ~ N(μ,σ²)

Y = aX + b

E[Y] = E[aX+b] = aμ + b

Var[Y] = Var[aX+b] = a²σ²

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

normal random variables

X is a normal random variable   X ~ N(μ,σ²)

Y = aX + b

E[Y] = E[aX+b] = aμ + b

Var[Y] = Var[aX+b] = a²σ²

Y ~ N(aμ + b, a²σ²)

E[ ],Var[ ] as expected;
"normality" is the surprise

Important special case: Z = (X-μ)/σ ~ N(0,1)

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

normal random variables

X is a normal random variable   X ~ N(μ,σ²)

Y = aX + b

E[Y] = E[aX+b] = aμ + b

Var[Y] = Var[aX+b] = a²σ²

Y ~ N(aμ + b, a²σ²)

E[ ],Var[ ] as expected;
"normality" is the surprise

Important special case: Z = (X-μ)/σ ~ N(0,1)

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

Z ~ N(0,1)   "*standard* (or *unit*) normal"

Use Φ(z) to denote CDF, i.e.

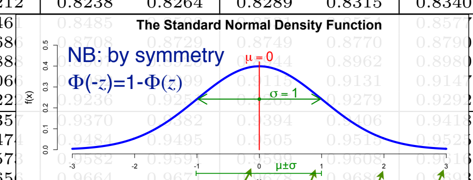
$$\Phi(z) = \Pr(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx$$

no closed form ☹

Table of the Standard Normal Cumulative Distribution Function Φ(z)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	[illegible]	0.7190
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	[illegible]	0.7517
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	[illegible]	0.7823
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365
1.0	0.8413	0.8438	0.8461	0.8485	[illegible]	[illegible]	[illegible]	0.8577	0.8599
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810
1.2	0.8849	0.8869	0.8888	[illegible]	[illegible]	[illegible]	0.8962	0.8980	0.8997
1.3	0.9032	0.9049	0.9066	[illegible]	[illegible]	0.9115	0.9131	0.9147	0.9162
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	[illegible]	0.9279	0.9292	0.9306
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	[illegible]	[illegible]	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	[illegible]	[illegible]	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997

Φ(.46)



E.g., see B&T p155, p531

X normal with mean 3 and variance 9.
What is

Pr (X > 0)

Pr (2 < X < 5)

Pr (|X-3| > 6)

Table of the Standard Normal Cumulative Distribution Function $\Phi(z)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997

X normal with mean 3 and variance 9.

$$Pr(2 < X < 5) = Pr(\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}) = Pr(-\frac{1}{3} < Z < \frac{2}{3})$$

$$= \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right) \quad = \Phi\left(\frac{2}{3}\right) - \left(1 - \Phi\left(\frac{1}{3}\right)\right)$$

$$= 0.7456 - (1 - 0.6293)$$

$$Pr(X > 0) = Pr(Z > \frac{0-3}{3}) = Pr(Z > -1) = Pr(Z < 1) = 0.8413$$

$$Pr(|X-3| > 6) = Pr(X > 9) + Pr(X < -3) =$$
$$Pr(Z > \frac{9-3}{3}) + Pr(Z < \frac{-3-3}{3})$$

Solution

X normal with mean 5 and variance σ^2
If Pr(X> 9) = 0.2, then approximately what is σ^2 ?

$$Pr(X > 9) = Pr(\frac{X-5}{\sigma} > \frac{9-5}{\sigma}) = 0.2$$

$$1 - \Phi\left(\frac{9-5}{\sigma}\right) = 0.2$$

$$\Phi\left(\frac{9-5}{\sigma}\right) = 0.8$$

Look up in N(0,1) table to find our what v gives $\Phi(v) = 0.8$

Set $\frac{9-5}{\sigma} = v$ and solve for σ

pdf vs cdf

$f(x) = \frac{d}{dx} F(x)$ $F(a) = \int_{-\infty}^{a} f(x)\, dx$

sums become integrals, e.g.

$E[X] = \Sigma_x\, xp(x)$ $E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$

most familiar properties still hold, e.g.

$E[aX+bY+c] = aE[X]+bE[Y]+c$

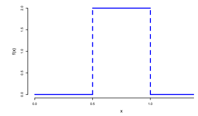
$Var[X] = E[X^2] - (E[X])^2$

Three important examples

$X \sim Uni(\alpha,\beta)$ uniform in $[\alpha,\beta]$

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha,\beta] \\ 0 & \text{otherwise} \end{cases}$$

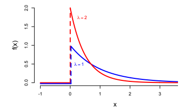
$E[X] = (\alpha+\beta)/2$
$Var[X] = (\alpha-\beta)^2/12$



$X \sim Exp(\lambda)$ exponential

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

$E[X] = \frac{1}{\lambda}$
$\mathrm{Var}[X] = \frac{1}{\lambda^2}$



$X \sim N(\mu, \sigma^2)$ normal (aka Gaussian, aka the big Kahuna)

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

$E[X] = \mu$
$\mathrm{Var}[X] = \sigma^2$