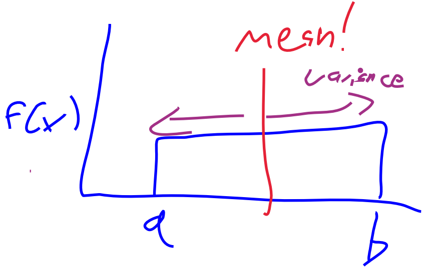
CSCI 3022-002 Intro to Data Science
The Normal Distribution



What are the mean and variance of the continuous uniform distribution?

Integrals!

Announcements and Reminders

- Practicum due Monday!

Exam the next one.

ALL SPORTS COMMENTARY

In practice, we often just look up the formulas for the pdfs, means, and variances of whatever model we choose to use.

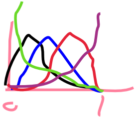
Discrete

Table of Common Distributions

taken from *Statistical Inference* by Casella and Berger

Discrete Distrbutions

distribution	pmf	mean	variance	mgf/moment
Bernoulli(p)	$p^x(1-p)^{1-x};\ x=0,1;\ p\in(0,1)$	p	$p(1-p)$	$(1-p)+pe^t$
Beta-binomial(n,α,β)	$\binom{n}{x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)}{\Gamma(\alpha+\beta+n)}$	$\frac{n\alpha}{\alpha+\beta}$	$\frac{n\alpha\beta}{(\alpha+\beta)^2}$	
Notes: If $X\|P$ is binomial (n,P) and P is beta(α,β), then X is beta-binomial(n,α,β).				
Binomial(n,p)	$\binom{n}{x}p^x(1-p)^{n-x};\ x=1,\ldots,n$	np	$np(1-p)$	$[(1-p)+pe^t]^n$
Discrete Uniform(N)	$\frac{1}{N};\ x=1,\ldots,N$	$\frac{N+1}{2}$	$\frac{(N+1)(N-1)}{12}$	$\frac{1}{N}\sum_{i=1}^{N}e^{it}$
Geometric(p)	$p(1-p)^{x-1};\ p\in(0,1)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Note: $Y=X-1$ is negative binomial$(1,p)$. The distribution is *memoryless*: $P(X>s\|X>t)=P(X>s-t)$.				
Hypergeometric(N,M,K)	$\frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}};\ x=1,\ldots,K$ $M-(N-K)\le x\le M;\ N,M,K>0$	$\frac{KM}{N}$	$\frac{KM}{N}\frac{(N-M)(N-k)}{N(N-1)}$	?
Negative Binomial(r,p)	$\binom{r+x-1}{x}p^r(1-p)^x;\ p\in(0,1)$ $\binom{y-1}{r-1}p^r(1-p)^{y-r};\ Y=X+r$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$	$\left(\frac{p}{1-(1-p)e^t}\right)^r$
Poisson(λ)	$\frac{e^{-\lambda}\lambda^x}{x!};\ \lambda\ge 0$	λ	λ	$e^{\lambda(e^t-1)}$
Notes: If Y is gamma(α,β), X is Poisson$(\frac{x}{\beta})$, and α is an integer, then $P(X\ge\alpha)=P(Y\le y)$.				



← Beta

Continuous Distributions

distribution	pdf	mean	variance	mgf/moment
Beta(α,β)	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1};\ x\in(0,1),\ \alpha,\beta>0$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}$
Cauchy(θ,σ)	$\frac{1}{\pi\sigma}\frac{1}{1+(\frac{x-\theta}{\sigma})^2};\ \sigma>0$	does not exist	does not exist	does not exist
Notes: Special case of Students's t with 1 degree of freedom. Also, if X,Y are iid $N(0,1)$, $\frac{X}{Y}$ is Cauchy				
χ^2_p	$\frac{1}{\Gamma(\frac{p}{2})2^{\frac{p}{2}}}x^{\frac{p}{2}-1}e^{-\frac{x}{2}};\ x>0,\ p\in N$	p	$2p$	$\left(\frac{1}{1-2t}\right)^{\frac{p}{2}},\ t<\frac{1}{2}$
Notes: Gamma$(\frac{p}{2},2)$.				
Double Exponential(μ,σ)	$\frac{1}{2\sigma}e^{-\frac{\lvert x-\mu\rvert}{\sigma}};\ \sigma>0$	μ	$2\sigma^2$	$\frac{e^{\mu t}}{1-(\sigma t)^2}$
Exponential(θ)	$\frac{1}{\theta}e^{-\frac{x}{\theta}};\ x\ge 0,\ \theta>0$	θ	θ^2	$\frac{1}{1-\theta t},\ t<\frac{1}{\theta}$
Notes: Gamma$(1,\theta)$. Memoryless. $Y=X^{\frac{1}{\gamma}}$ is *Weibull.* $Y=\sqrt{\frac{2X}{\beta}}$ is *Rayleigh.* $Y=\alpha-\gamma\log\frac{X}{\beta}$ is *Gumbel.*				
F_{ν_1,ν_2}	$\frac{\Gamma(\frac{\nu_1+\nu_2}{2})}{\Gamma(\frac{\nu_1}{2})\Gamma(\frac{\nu_2}{2})}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}\frac{x^{\frac{\nu_1-2}{2}}}{\left(1+(\frac{\nu_1}{\nu_2})x\right)^{\frac{\nu_1+\nu_2}{2}}};\ x>0$	$\frac{\nu_2}{\nu_2-2},\ \nu_2>2$	$2(\frac{\nu_2}{\nu_2-2})^2\frac{\nu_1+\nu_2-2}{\nu_1(\nu_2-4)},\ \nu_2>4$	$EX^n=\frac{\Gamma(\frac{\nu_1+2n}{2})\Gamma(\frac{\nu_2-2n}{2})}{\Gamma(\frac{\nu_1}{2})\Gamma(\frac{\nu_2}{2})}\left(\frac{\nu_2}{\nu_1}\right)^n,\ n<\frac{\nu_2}{2}$
Notes: $F_{\nu_1,\nu_2}=\frac{\chi^2_{\nu_1}/\nu_1}{\chi^2_{\nu_2}/\nu_2}$, where the χ^2s are independent. $F_{1,\nu}=t^2_\nu$.				
Gamma(α,β)	$\frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-\frac{x}{\beta}};\ x>0,\ \alpha,\beta>0$	$\alpha\beta$	$\alpha\beta^2$	$\left(\frac{1}{1-\beta t}\right)^{\alpha},\ t<\frac{1}{\beta}$
Notes: Some special cases are exponential $(\alpha=1)$ and χ^2 $(\alpha=\frac{p}{2},\beta=2)$. If $\alpha=\frac{2}{3}$, $Y=\sqrt{\frac{X}{\beta}}$ is *Maxwell.* $Y=\frac{1}{X}$ is *inverted gamma.*				
Logistic(μ,β)	$\frac{1}{\beta}\frac{e^{-\frac{x-\mu}{\beta}}}{\left[1+e^{-\frac{x-\mu}{\beta}}\right]^2};\ \beta>0$	μ	$\frac{\pi^2\beta^2}{3}$	$e^{\mu t}\Gamma(1+\beta t),\ \lvert t\rvert<\frac{1}{\beta}$
Notes: The cdf is $F(x\mid\mu,\beta)=\frac{1}{1+e^{-\frac{x-\mu}{\beta}}}$.				
Lognormal(μ,σ^2)	$\frac{1}{\sqrt{2\pi}\sigma}\frac{1}{x}e^{-\frac{(\log x-\mu)^2}{2\sigma^2}};\ x>0,\sigma>0$	$e^{\mu+\frac{\sigma^2}{2}}$	$e^{2(\mu+\sigma^2)}-e^{2\mu+\sigma^2}$	$EX^n=e^{n\mu+\frac{n^2\sigma^2}{2}}$
Normal(μ,σ^2)	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}};\ \sigma>0$	μ	σ^2	$e^{\mu t+\frac{\sigma^2t^2}{2}}$
Pareto(α,β)	$\frac{\beta\alpha^\beta}{x^{\beta+1}};\ x>\alpha,\ \alpha,\beta>0$	$\frac{\beta\alpha}{\beta-1},\ \beta>1$	$\frac{\beta\alpha^2}{(\beta-1)^2(\beta-2)},\ \beta>2$	does not exist
t_ν	$\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}\frac{1}{\sqrt{\nu\pi}}\frac{1}{(1+\frac{x^2}{\nu})^{\frac{\nu+1}{2}}}$	$0,\ \nu>1$	$\frac{\nu}{\nu-2},\nu>2$	$EX^n=\frac{\Gamma(\frac{\nu+1}{2})\Gamma(\nu-\frac{n}{2})}{\sqrt{\pi}\Gamma(\frac{\nu}{2})}\nu^{\frac{n}{2}},\ n$ even
Notes: $t^2_\nu=F_{1,\nu}$.				
Uniform(a,b)	$\frac{1}{b-a},\ a\le x\le b$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{(b-a)t}$
Notes: If $a=0,\ b=1$, this is special case of beta $(\alpha=\beta=1)$.				
Weibull(γ,β)	$\frac{\gamma}{\beta}x^{\gamma-1}e^{-\frac{x^\gamma}{\beta}};\ x>0,\ \gamma,\beta>0$	$\beta^{\frac{1}{\gamma}}\Gamma(1+\frac{1}{\gamma})$	$\beta^{\frac{2}{\gamma}}\left[\Gamma(1+\frac{2}{\gamma})-\Gamma^2(1+\frac{1}{\gamma})\right]$	$EX^n=\beta^{\frac{n}{\gamma}}\Gamma(1+\frac{n}{\gamma})$
Notes: The mgf only exists for $\gamma\ge 1$.				

t or

sum of exponentials

Variance and EV Recap

1. **Expected Value:** The average value for X coming from a distribution (not a sample!). Denoted $E[X]$ or μ or μ_X.
   Discrete: $\sum_{x \in \Omega} x f(x)$; Continuous: $\int_{x \in \Omega} x \cdot f(x)\, dx$
2. Expected value of a function $g(X)$ of X is:
   $\sum_{x \in \Omega} g(x) f(x)$; $\int_{x \in \Omega} g(x) \cdot f(x)\, dx$
3. $Y = g(X)$ is a *change of variables*.
4. Expectation is **linear:** $E[aX + b] = aE[X] + b$
5. Variance isn't linear: $Var[aX + b] = a^2 Var[X]$
6. $Var[X] = E[X^2] - E[X]^2$ $=$ $E[(X - E[X])^2]$

Known Variances and EVs

Dist.	Mean	Variance
Bernoulli	p	$p(1-p)$
Binomial	np	$np(1-p)$
Geometric	$1/p$	$\frac{p^2}{1-p}$
Poisson	λ	λ
Cont. Unif	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential	$1/\lambda$	$1/\lambda^2$

A Variance shortcut

$E[X] = \int x f(x)\,dx$

$E[X^2] = \int x^2 f(x)\,dx$

easier than

$E[(X-\mu_X)^2] = \int (x - E[X])^2 f(x)\,dx$

note: for a r.v. X, $E[X]$ is not random.

$E[4] = 4 \Rightarrow E[E[X]] = E[X]$

When computing variance, it's often easier to use the following formula:

$$\underbrace{E\left[(X - E(X))^2\right]}_{Var[X]} = \boxed{E[X^2] - E[X]^2} \quad \text{soln!}$$

Proof:

$$(\;)= E[X^2 - 2XE[X] + (E[X])^2]$$

$$= E[X^2] - 2E[E[X] \cdot X] + E[(E[X])^2]$$

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

Opening sol:

The pdf is $f(x) = \frac{1}{b-a}$ in $[a, b]$

mean: $E[x] = \int_a^b x \cdot f(x)\,dx = \int_a^b \frac{1}{b-a}\, x \, dx$

$= \frac{1}{b-a} \int_a^b x\,dx$ scalar

$\frac{1}{b-a}\left(\frac{x^2}{2}\right)\Big|_a^b = \frac{1}{b-a}\left(\frac{b^2-a^2}{2}\right)$

$= \cancel{\frac{1}{b-a}\left(\frac{(b-a)(b+a)}{2}\right)} = \frac{b+a}{2}$

$E[X^2] = \int_a^b \frac{1}{b-a}\, x^2\,dx = \frac{x^3}{3}\Big|_a^b \cdot \frac{1}{(b-a)}$

$= \frac{1}{b-a}\left(\frac{b^3-a^3}{3}\right)$

then use

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

Opening sol:

The pdf is $f(x) = \frac{1}{b-a}$ in $[a, b]$
It's on the prior slide's tables. Nailed it!

Opening sol:

The pdf is $f(x) = \frac{1}{b-a}$ in $[a, b]$

The mean is $\int_a^b \frac{1}{b-a} x\, dx$, so

$$E[X] = \frac{1}{b-a}\frac{x^2}{2}\Big|_a^b = \frac{1}{b-a}\frac{b^2-a^2}{2} = \frac{(b-a)(b+a)}{2} = \frac{a+b}{2}$$

Opening sol:

The pdf is $f(x) = \frac{1}{b-a}$ in $[a, b]$

The mean is $\int_a^b \frac{1}{b-a} x\, dx$, so

$$E[X] = \frac{1}{b-a}\frac{x^2}{2}\Big|_a^b = \frac{1}{b-a}\frac{b^2-a^2}{2} = \frac{(b-a)(b+a)}{2} = \frac{a+b}{2}$$

The variance is probably easier to compute using the shortcut formula. So lets find

$$E[X^2] = \int_a^b \frac{1}{b-a} x^2\, dx = \frac{1}{b-a}\frac{x^3}{3}\Big|_a^b$$

$$= \frac{1}{b-a}\frac{b^3-a^3}{3} = \frac{(b-a)(b+a)}{3} = \frac{a^2+ab+b^2}{3}$$

Combining this with the mean squared, we have:

$$Var[X] = E[X^2] - E[X]^2 = \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4}$$

$$= \frac{a^2-2ab+b^2}{12} = \frac{(b-a)^2}{12}$$

X: outcomes: # sold	$P(X=x)$	Profit	(sold · 500 + unsold (-300))
0	.1	-900	Profit^2
1	.2	-100	
2	.3	700	
3	.4	1500	

A computer store has purchased 3 computers of a certain type at \$500 each. It will sell them for \$1000 each. The manufacturer has agreed to repurchase any computers still unsold after a specified period at \$200 each. Let X denote the number of computers sold, and suppose that:

$P(X=0)=0.1, P(X=1)=0.2, P(X=2)=0.3, P(X=3)=0.4.$

What is the expected profit? What is the standard deviation of the profit?

$E[\text{profit}] = E[X\,500 + (3-X)\cdot(-300)]$

Sum outcome · its probability!

$.1(-900) + .2(-100) + .3(700) + .4(1500) = -90 - 20 + 210 + 600 = \$700!$

Var

$E[\text{profit}^2]$

$.1\cdot(-900)^2 + .2(-100)^2 + .3(700)^2 + .4(1500)^2$

Then: $E[\text{profit}^2] - E[\text{profit}]^2$

The Normal Distribution

The normal distribution (sometimes called the Gaussian distribution) is probably the most important ~~distribution in all of probability and statistics~~.

function in the world.

Many populations have distributions that can be fit very closely by an appropriate normal (or Gaussian, bell) curve.

Examples: height, weight, and other physical characteristics, scores on various tests, etc.

The Normal Distribution

Definition: *Normal Distribution:*

A continuous r.v. X is said to have a *normal distribution* with parameters μ and $\sigma > 0$, if the pdf of X is:

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

sigma: width/dispersion

st. dev.

Notation: We write $N(\mu, \sigma^2)$

mean

variance

mean: center

The Normal Distribution

Definition: *Normal Distribution:*

A continuous r.v. X is said to have a *normal distribution* with parameters $\underline{\mu}$ and $\underline{\sigma^2} > 0$, if the pdf of X is:

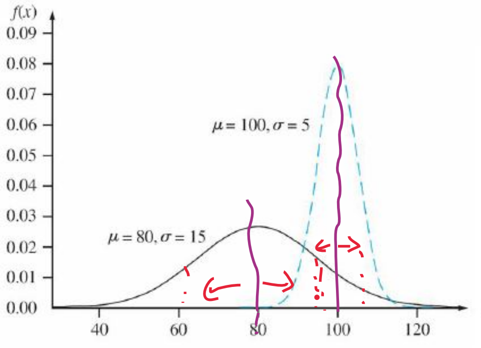
$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2\sigma^2}(x-\mu)^2}$$

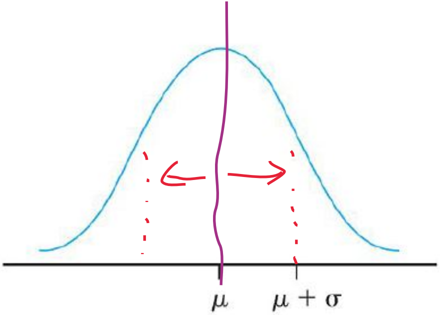
Notation: We write $\underline{X \sim N(\mu\sigma^2)}$

The Normal Distribution

The figure below presents graphs of f for different parameter pairs:

Center: μ





You can play with normals in any statistical software. See for example
`https://academo.org/demos/gaussian-distribution/`

The Standard Normal Distribution

Definition: *Standard Normal Distribution:*
The normal distribution with parameter values $\mu = 0$ and $\sigma = 1$ is called the *standard normal distribution*.

A r.v. with this distribution is called a standard normal random variable and is denoted by $\boxed{Z}$.
Its pdf is:

$$f(z) = \frac{1}{\sqrt{2\pi}\,\Box} \cdot e^{-\frac{1}{2\Box}(x\ \Box)^2}$$

$\mu_0 \sim 0$

$\sigma \cdot 1$

The Standard Normal Distribution

Definition: *Standard Normal Distribution:*
The normal distribution with parameter values $\underline{\mu = 0}$ and $\underline{\sigma^2 = 1}$ is called the *standard normal distribution*.

A r.v. with this distribution is called a standard normal random variable and is denoted by Z. Its pdf is:

$$f(z) =$$

The Standard Normal Distribution

Definition: *Standard Normal Distribution:*
The normal distribution with parameter values $\underline{\mu = 0}$ and $\underline{\sigma^2 = 1}$ is called the *standard normal distribution*.

A r.v. with this distribution is called a standard normal random variable and is denoted by Z. Its pdf is:

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

The normal cdf

Let's find the cdf of the standard normal distribution!
All we have to to is integrate:

$$\int_{-\infty}^{Z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$$

The normal cdf

Let's find the cdf of the standard normal distribution!
All we have to to is integrate:

Phi!

$$\int_{-\infty}^{Z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt = F(z) = \Phi(z)$$

Should we try a substitution? IBP?... this may not go integreat for us.

The normal cdf

Let's find the cdf of the standard normal distribution!
All we have to to is integrate:

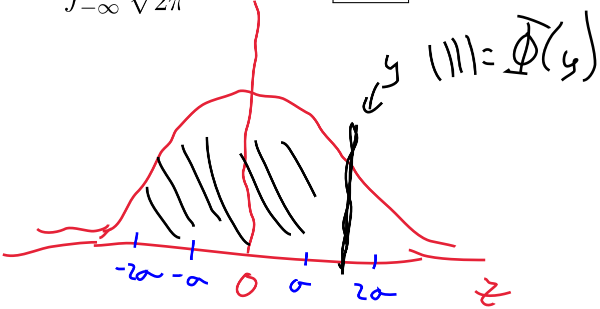
$$\int_{-\infty}^{Z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$$

The CDF of the normal distribution has no closed form. But it's really important! So we give it it's own name.

The normal cdf

For a random variable $Z \sim N(0,1)$, the cdf of Z is given by

$$F(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt = \boxed{\Phi(z)}$$



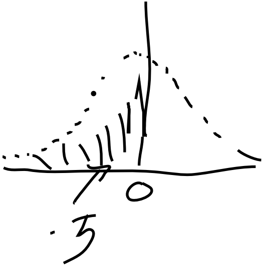
The normal cdf

For a random variable $Z \sim N(0,1)$, the cdf of Z is given by

$$F(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt = \boxed{\Phi(z)}$$

Old school statisticians used to carry around giant tables with values of $\Phi(z)$ in them. Actually, many current statisticians do that too, but that's a little silly. We have computers!

The Standard Normal



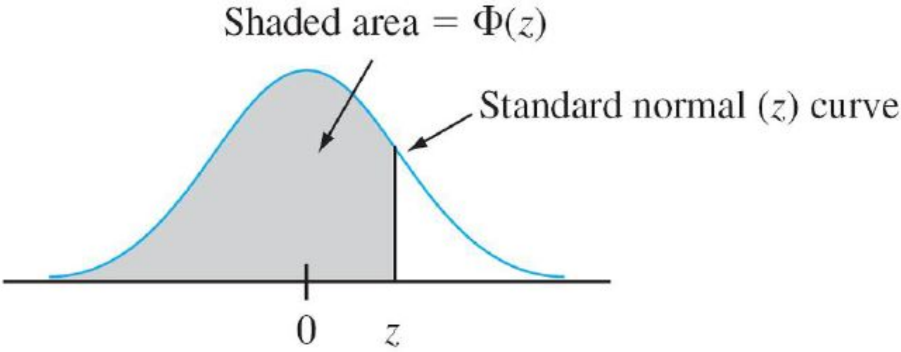
Note:

1. The standard normal distribution rarely occurs naturally.
2. Instead, it is a reference distribution from which information about other normal distributions can be obtained via a simple formula.
3. These probabilities can then be found "normal tables".
4. This can also be computed with a single command... (`scipy.stats.norm.cdf`, for example)

scipy.stats.norm.cdf(0) = .5

The Standard Distribution

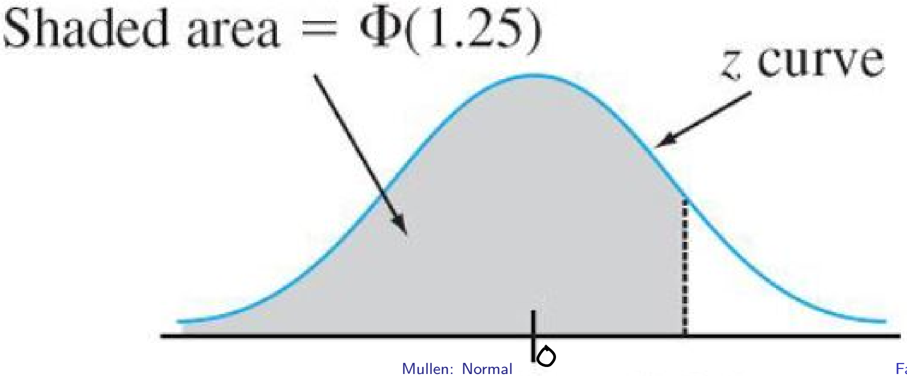
The figure below illustrates the probabilities found in a normal table (such a table can easily be found online):



The Standard Distribution

$P(Z \leq 1.25) = \Phi(1.25)$, a probability that is tabulated in a normal table. What is this probability?

The figure below illustrates this probability:



The Standard Distribution

we can do: $\Phi(z) = P(Z \leq z)$

$:=$ stats.norm.cdf(z)

Some quick examples:

1. $P(Z \geq 1.25)$ $= 1 - P(Z < 1.25) = 1 - P(Z \leq 1.25)$

   $= 1 - \Phi(1.25)$

2. Why does $P(Z < -1.25) = P(Z > 1.25)$? What is $\Phi(-1.25)$?

3. How do we calculate $P(-.38 \leq Z \leq 1.25)$?

   $= \int_{-.38}^{1.25} f(z)\, dz$

   $= \int_{-\infty}^{1.25} f(z)dz + \boxed{\int_{-.38}^{-\infty} f}$

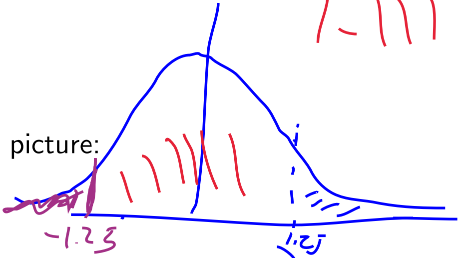
   $= \int_{-\infty}^{-.38} f(z)dz$

   we know $\int_{-\infty}^{z} f(x)dx$

The Standard Distribution

Some quick examples:

1. $P(Z \geq 1.25) =$ $P(Z \leq -1.25)$
   It's `1-scipy.stats.norm.cdf(1.25)`. Or as a picture:



2. Why does $P(Z < -1.25) = P(Z > 1.25)$? What is $\Phi(-1.25)$?
   Symmetry! Same as above.

3. How do we calculate $P(-.38 \leq Z \leq 1.25)$?
   As an integral, this is $\int_{-.38}^{1.25} f(z)\,dz$. We could split this into 2:
   $$\int_{-\infty}^{1.25} f(z)\,dz + \int_{-.38}^{-\infty} f(z)\,dz =$$
   $$\Phi(1.25) - \Phi(-.38)$$

Standard Quantiles

The 99th *percentile* of the standard normal distribution is that value of z such that the area under the z curve to the left of the value is 0.99.

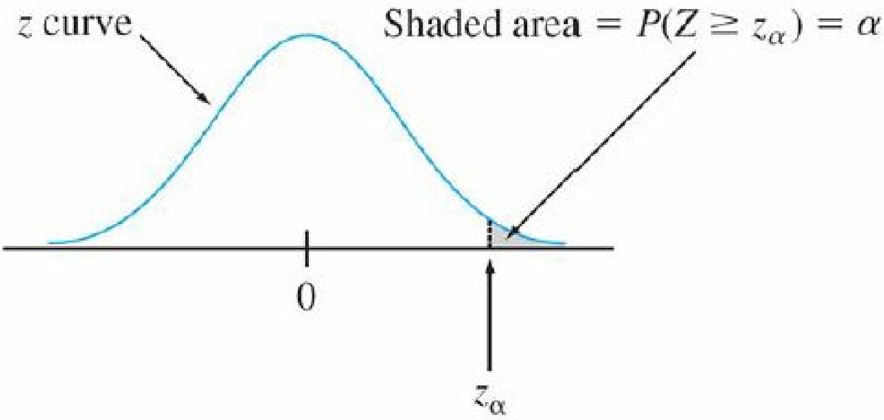
Tables and cdf functions give, for fixed z, the area under the standard normal curve to the left of z; now we have the area and want the value of z.

This is the "inverse" problem to $P(Z \leq z) = ?$

How can the table be used for this?

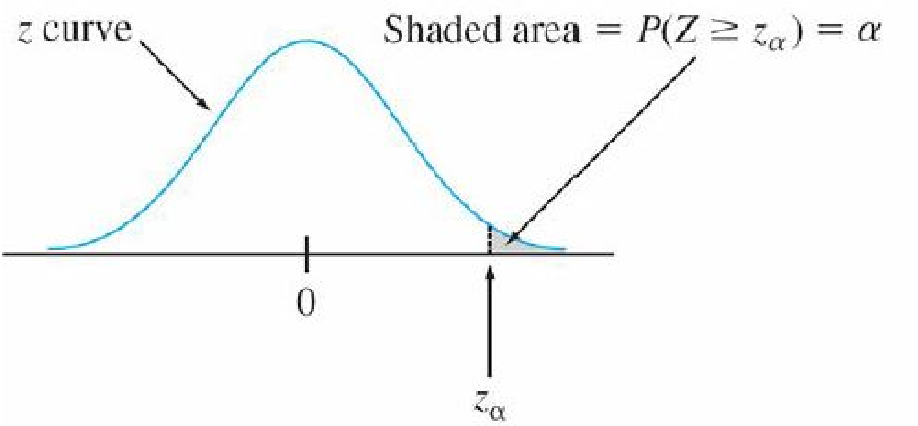
Standard Quantiles

In statistical inference, we need the z values that give certain tail areas under the standard normal curve. There, this notation will be standard: __ will denote the z value for which _ of the area under the z curve lies to the right of __.



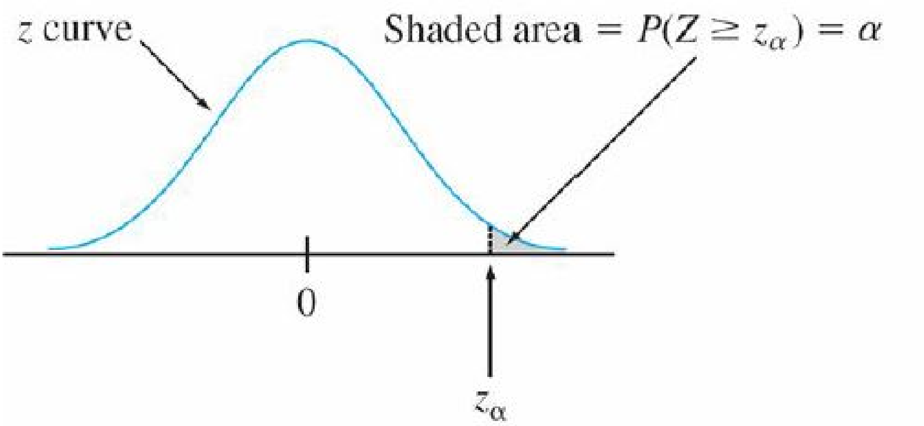
Standard Quantiles

In statistical inference, we need the z values that give certain tail areas under the standard normal curve. There, this notation will be standard: $\underline{z_\alpha}$ will denote the z value for which $\underline{\alpha}$ of the area under the z curve lies to the right of $\underline{z_\alpha}$.



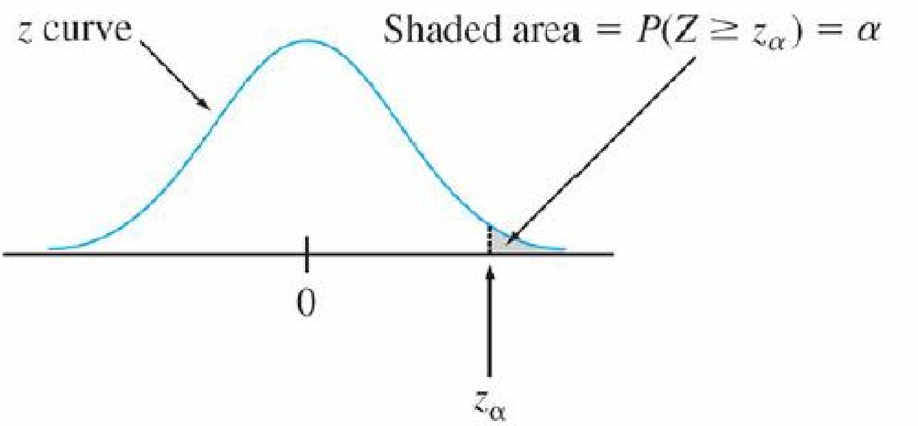
Standard Quantiles

In statistical inference, we need the z values that give certain tail areas under the standard normal curve. There, this notation will be standard: __ will denote the z value for which _ of the area under the z curve lies to the right of __.



Standard Quantiles

In statistical inference, we need the z values that give certain tail areas under the standard normal curve. There, this notation will be standard: $\underline{z_\alpha}$ will denote the z value for which $\underline{\alpha}$ of the area under the z curve lies to the right of $\underline{z_\alpha}$.



Non-Standard Normals

When $\underline{X \sim N(\mu, \sigma^2)}$, probabilities involving X are computed by "standardizing." The standardized variable is:

Proposition: If X has a normal distribution with mean _ and standard deviation _, then

is distributed standard normal.

Non-Standard Normals

When $\underline{X \sim N(\mu, \sigma^2)}$, probabilities involving X are computed by "standardizing." The standardized variable is:

random minus its mean

$$Z = \frac{X - \mu}{\sigma}$$

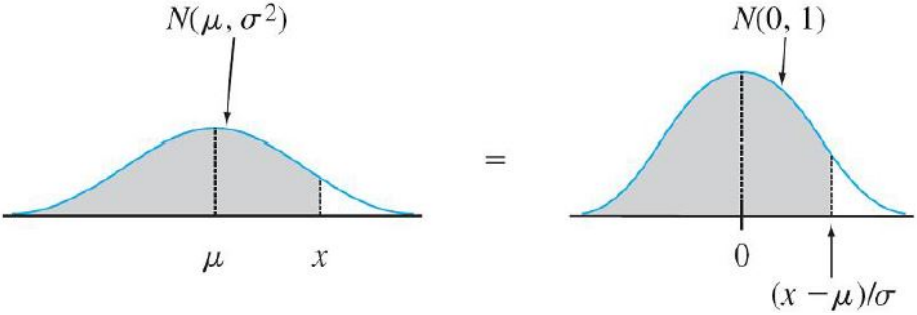
divided its st. dev.

Proposition: If X has a normal distribution with mean $\underline{\mu}$ and standard deviation $\underline{\sigma}$, then

is distributed standard normal.

Non-Standard Normals

Why do we standardize normal random variables?



Equality of nonstandard and standard normal curve areas

Using Normals

Example:
The time that it takes a driver to react to the brake lights on a decelerating vehicle is critical in helping to avoid rear-end collisions.

Research suggests that reaction time for an in-traffic response to a brake signal from standard brake lights can be modeled with a normal distribution having mean value 1.25 sec and standard deviation of 0.46 sec.

What is the probability that reaction time is between 1.00 sec and 1.75 sec?

Solution:

Example: For a normal distribution having mean value 1.25 sec and standard deviation of 0.46 sec.

What is the probability that reaction time is between 1.00 sec and 1.75 sec?

Solution:

Example: For a normal distribution having mean value 1.25 sec and standard deviation of 0.46 sec.

$$X \sim N(1.25, .46)$$

What is the probability that reaction time is between 1.00 sec and 1.75 sec?
We want $P(1 < X < 1.75)$... but we can't compute these probabilities unless the r.v. in the middle of the inequality is *standard* normal. So we normalize!

Solution:

Example: For a normal distribution having mean value 1.25 sec and standard deviation of 0.46 sec.

What is the probability that reaction time is between 1.00 sec and 1.75 sec?
We want $P(1 < X < 1.75)$... but we can't compute these probabilities unless the r.v. in the middle of the inequality is *standard* normal. So we normalize!

$$P(1 < X < 1.75) = P(1 - 1.25 < X - 1.25 < 1.75 - 1.25)$$

$$= P(\frac{-.25}{.46} < \frac{X - 1.25}{.46} < \frac{.5}{.46}) = P(\frac{-.25}{.46} < Z < \frac{.5}{.46})$$

$$= \Phi(\frac{-.25}{.46}) - \Phi(\frac{.5}{.46})$$

Daily Recap

Today we learned

1. The Normal!

Moving forward:

- nb day Friday

Next time in lecture:

- Beginning: why we care so much about the normal!