Confidence intervals for the population variance σ^2 based on the sample variance s^2

Confidence interval for the population variance σ^2

- Up until now we were calculating the confidence interval on the **population average μ**
- What if one wants to put **confidence interval on the population variance $\boldsymbol{\sigma^2}$**?
- We know an unbiased estimator of σ^2:

$$s^2 = \frac{1}{n-1}\sum_i (x_i - \bar{x})^2$$

- How to determine the confidence interval?

8-4 Confidence Interval on the Variance and Standard Deviation of a Normal Distribution

Definition

(Eq. 8-17)

Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution with mean μ and variance σ^2, and let S^2 be the sample variance. Then the random variable

$$X^2 = \frac{(n-1)S^2}{\sigma^2} \qquad (8\text{-}17)$$

has a chi-square (χ^2) distribution with $n-1$ degrees of freedom.

8-4 Confidence Interval on the Variance and Standard Deviation of a Normal Distribution

$X=(n-1)S^2/\sigma^2$

We know n, S^2

want to estimate σ^2

$f(x,n) \sim x^{(n-1)/2-1}\exp(-x/2)$

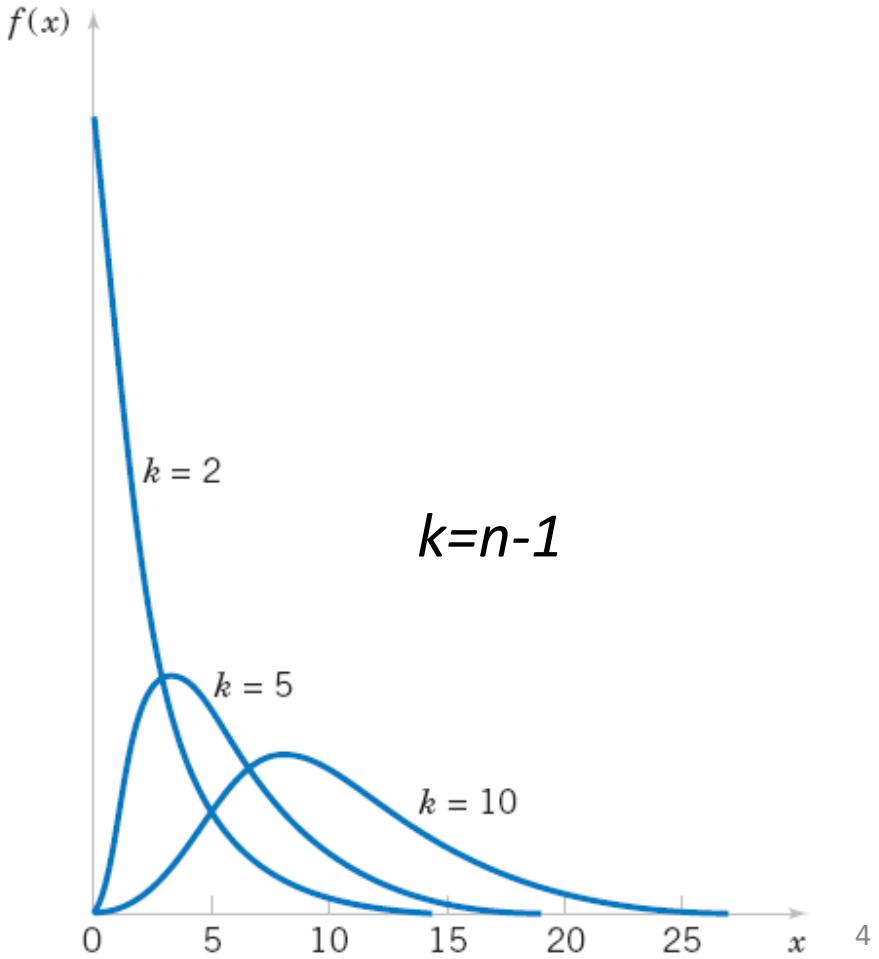
It is just Gamma PDF with $r=(n-1)/2$, and $\lambda=1/2$

Mean value:

$n-1$

Standard deviation:

$$\sqrt{2(n-1)}$$



Number of degrees of freedom $k = n - 1$

Let $X = (X_1, \dots, X_n)$ iid $\mathcal{N}(0,1)$, $\bar{X} = \frac{1}{n}\sum X_i$, and consider the **residual sum of squares**

$$q = \sum_{i=1}^{n} (X_i - \bar{X})^2 .$$

Geometrically, decompose $\mathbb{R}^n$ into the span of $u = \frac{1}{\sqrt{n}}(1, \dots, 1)$ and its orthogonal complement $u^{\perp}$ (dimension $n - 1$). There exists an orthogonal matrix U with first column u, so

$$Y = U^{\top} X \sim \mathcal{N}_n(0, I_n), \quad Y_1 = \langle X, u \rangle = \sqrt{n}\,\bar{X}, \quad (Y_2, \dots, Y_n) \in u^{\perp}.$$

$$q = \sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{j=2}^{n} Y_j^2 = \| Y_{\perp} \|^2,$$

where $Y_{\perp} = (Y_2, \dots, Y_n) \sim \mathcal{N}_{n-1}(0, I_{n-1})$ and is **independent** of Y_1.

Now integrate in the $(n-1)$-dimensional space $u^{\perp}$. In $\mathbb{R}^{n-1}$, the surface area of the $(n-2)$-sphere of radius r is

$$S_{n-2}(r) = \frac{2\pi^{(n-1)/2}}{\Gamma(n - 1/2)}\, r^{n-2}.$$

The radial density for $R = \| Y_{\perp} \|$ is

$$f_R(r) = \frac{S_{n-2}(r)}{(2\pi)^{(n-1)/2}}\, e^{-r^2/2} = \frac{2^{1-(n-1)/2}}{\Gamma(n - 1/2)}\, r^{n-2} e^{-r^2/2}, \quad r > 0.$$

With $q = r^2$ and $dr = \frac{1}{2\sqrt{q}}\, dq$,

$$f_Q(q) = f_R(\sqrt{q})\, \frac{1}{2\sqrt{q}} = \frac{1}{2^{(n-1)/2}\Gamma(n - 1/2)}\, q^{\frac{n-1}{2} - 1} e^{-q/2}, \quad q > 0.$$

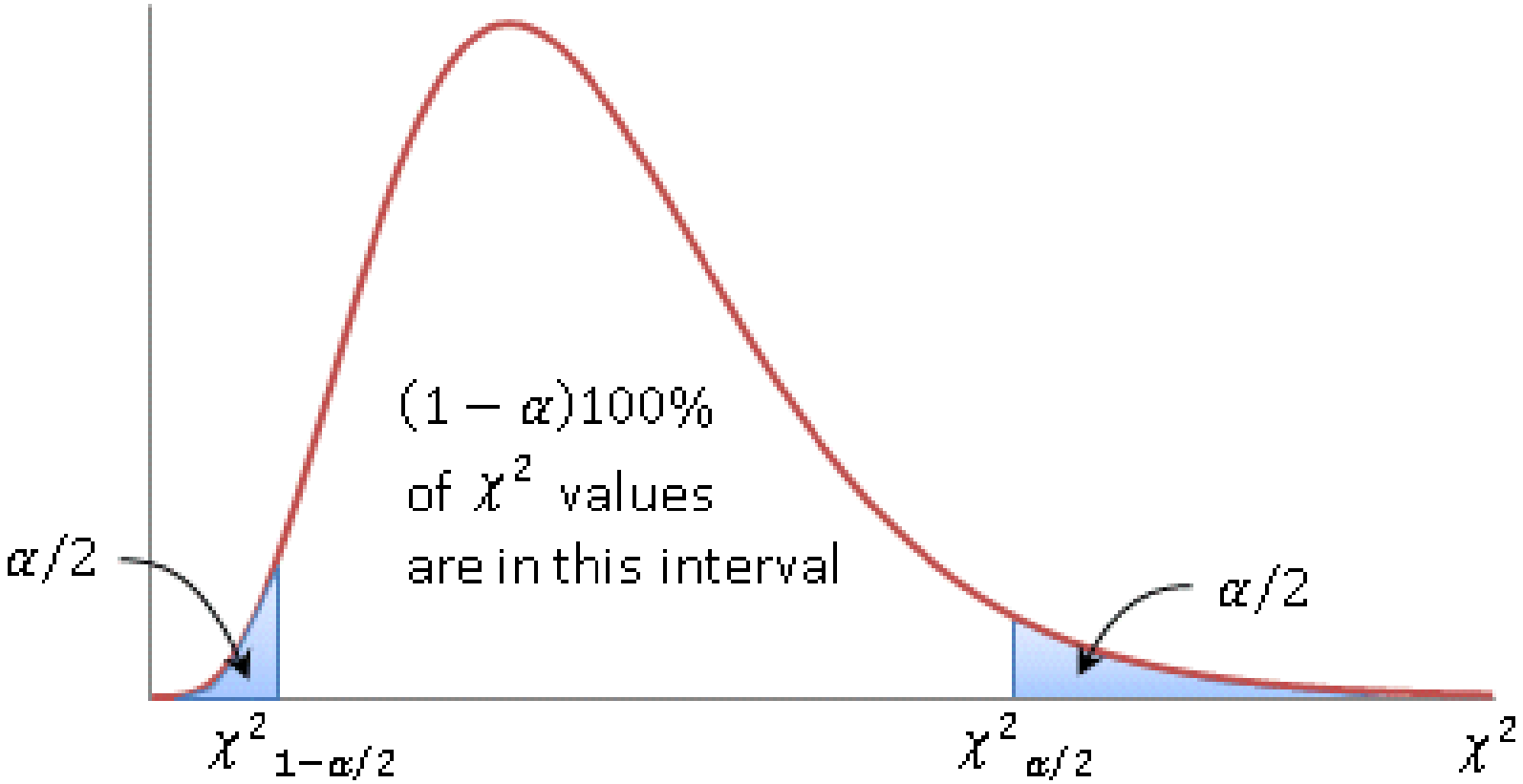
That is exactly χ^2_k with $k = n - 1$:

$$Q \sim \chi^2_{n-1}, \qquad f(q) = \frac{1}{2^{k/2}\Gamma(k/2)}\, q^{k/2-1} e^{-q/2}, \; k = n - 1.$$

Play with Mathematica notebook

http://demonstrations.wolfram.com/ChiSquaredDistributionAndTheCentralLimitTheorem/

By Peter Falloon



$$\chi^2_{1-\alpha/2} < \frac{(n-1)s^2}{\sigma^2} < \chi^2_{\alpha/2}$$

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}$$

8-4 Confidence Interval on the Variance and Standard Deviation of a Normal Distribution

Definition

(Eq. 8-19)

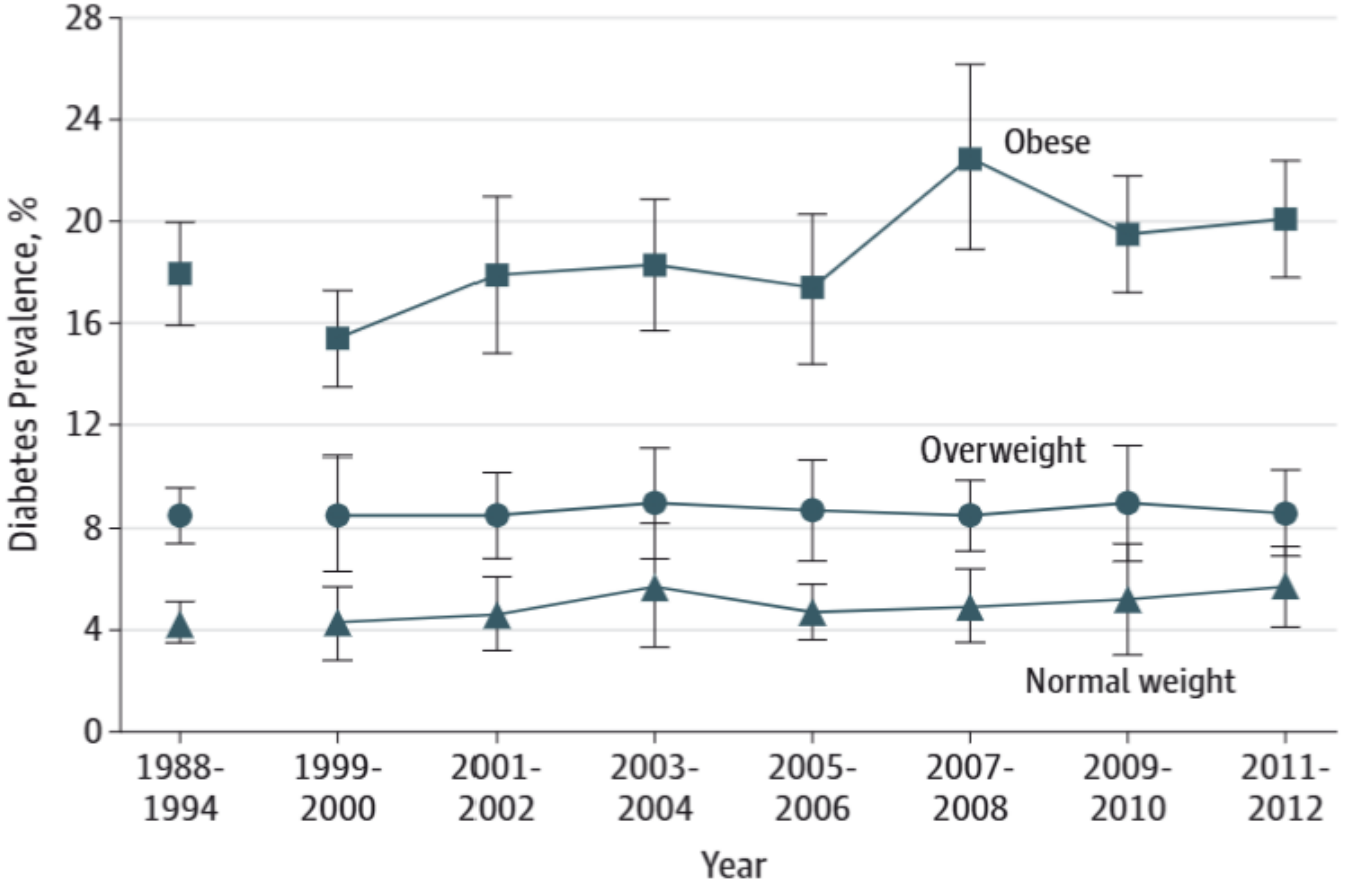
If s^2 is the sample variance from a random sample of n observations from a normal distribution with unknown variance σ^2, then **a $100(1-\alpha)\%$ confidence interval on σ^2 is**

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}} \qquad (8\text{-}19)$$

where $\chi^2_{\alpha/2,n-1}$ and $\chi^2_{1-\alpha/2,n-1}$ are the upper and lower $100\alpha/2$ percentage points of the chi-square distribution with $n-1$ degrees of freedom, respectively. A **confidence interval for σ** has lower and upper limits that are the square roots of the corresponding limits in Equation 8-19.

Confidence estimates of the population proportion

Figure 2. US Trends in Diabetes Prevalence per 100 Adults Aged 20 Years or Older by BMI Category



No. of participants								
Obese	2324	727	732	815	820	1137	1302	1075
Overweight	2942	724	878	784	694	949	1009	852
Normal weight	3025	645	699	624	604	726	762	785

Menke A, Casagrande S, Geiss L, Cowie CC JAMA. 2015;314: 1021–1029

Large sample confidence estimate of population proportion

- Want to know the fraction *p* of the population that belongs to a class, e.g., the class "people with diabetes"
- Each variable is a Bernoulli trial with one parameter p. We can use moments or MLE estimator to estimate p
- Both give the same estimate: sample fraction $\hat{p}$=*(# of people with diabetes in the sample)/(sample size n)*
- How to put confidence bounds on *p* based on $\hat{p}$
- Each participants in the sample is a Bernoulli trial: "success" = sampled participant has diabetes :-(
- Standard deviation of # of successes in n Bernoulli trials is given by Binomial distribution $\sqrt{n\,\hat{p}(1-\hat{p})}$
- Standard error of the fraction of successes is $\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$

8-5 A Large-Sample Confidence Interval For a Population Proportion (Eq. 8-23)

If $\hat{p}$ is the proportion of observations in a random sample of size n that belongs to a class of interest, an approximate $100(1-\alpha)\%$ confidence interval on the proportion p of the population that belongs to this class is

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \qquad (8\text{-}23)$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percentage point of the standard normal distribution.

This interval is known as the Wald interval (Wald and Wolfowitz, 1939).

Did you know that M&M's® Milk Chocolate Candies are supposed to come in the following percentages: 24% blue, 20% orange, 16% green, 14% yellow, 13% red, 13% brown?

http://www.scientificameriken.com/candy5.asp

"To our surprise M&Ms met our demand to review their procedures in determining candy ratios. It is, however, noted that the figures presented in their email differ from the information provided from their website (http://us.mms.com/us/about/products/milkchocolate/). An email was sent back informing them of this fact. To which M&Ms corrected themselves with one last email:

In response to your email regarding M&M'S CHOCOLATE CANDIES

Thank you for your email.
On average, our new mix of colors for M&M'S® Chocolate Candies is:

M&M'S® Milk Chocolate: 24% blue, 20% orange, 16% green, 14% yellow, 13% red, 13% brown.

M&M'S® Peanut: 23% blue, 23% orange, 15% green, 15% yellow, 12% red, 12% brown.

M&M'S® Kids MINIS®: 25% blue, 25% orange, 12% green, 13% yellow, 12% red, 13% brown.

M&M'S® Crispy: 17% blue, 16% orange, 16% green, 17% yellow, 17% red, 17% brown.

M&M'S® Peanut Butter and Almond: 20% blue, 20% orange, 20% green, 20% yellow, 10% red, 10% brown.



Have a great day!

Your Friends at Masterfoods USA
A Division of Mars, Incorporated

How to estimate these probabilities from a finite sample and how to set confidence interval on these estimates?

Did you know that M&M's® Milk Chocolate Candies are supposed to come in the following percentages: 24% blue, 20% orange, 16% green, 14% yellow, 13% red, 13% brown?



How large is a sample needed for 95% CI
on the percentage of blue M&Ms to be less than +/- 4%

Same question for red M&Ms?

Did you know that M&M's® Milk Chocolate Candies are supposed to come in the following percentages: 24% blue, 20% orange, 16% green, 14% yellow, 13% red, 13% brown?



How large is a sample needed for 95% CI
on the percentage of blue M&Ms to be less than +/- 4%

Same question for red M&Ms?

Hypothesis testing: one sample

Is P53 gene expressed at a lower level in cancer patients than in healthy people?

- We are interested if a P53 gene expression is **lowered** in population of cancer patients compared to the healthy population.
- We know that mean gene expression in the healthy population is μ_h=50 mRNAs/cell We are interested in deciding whether or not the mean expression in cancer population is **lower than** in healthy population. Let's call hypothesis H_1. Here H_1 is **one-sided**
- If we asked: cancer is not equal to healthy H_1 would be a **two-sided hypothesis**
- Assume we have a sample of 100 cancer patients with sample mean $\bar{x}$ =48 mRNAs/cell and standard deviation σ=10 mRNA/cell
- Can we use our sample to reject the "business as usual" or **null hypothesis H_0:** cancer = healthy and select **one-sided hypothesis H_1:** cancer < healthy

Two types of errors

	decide H_0	decide H_1
true H_0 probability	Correct action $1-\alpha$	Type I error α
true H_1 probability	Type II error β	Correct action power $= 1-\beta$

$$\alpha = P(\text{type I error}) = P(\text{reject } H_0 \text{ when } H_0 \text{ is true})$$

Sometimes the **type I error probability cutoff α** is called the **significance level,** or the **α-error cutoff**

Instructions: get the value of α from your boss or PI (e.g., 5% or 1%)

Prob(H_0 is true given the sample data) < α → reject H_0 and accept H_1

Prob(H_0 is true given the sample data) > α → accept H_0 and reject H_1

Type II error is much harder to estimate. We will deal with it later

P-Values of Hypothesis Tests

- P-value: what is the probability to get the observed value of sample mean of $\bar{x}$ =48 mRNAs/cell (or even smaller) and σ=10 mRNAs/cell in a healthy population with μ_h=50 mRNAs/cell

- If P-value is small – the null hypothesis is likely wrong and thus, the probability of making a type I error (incorrectly rejecting the null hypothesis) is small

- P-value answers the question: if I reject the null hypothesis H_0 based on the sample, what is the probability that I am making a type I error?

P-Value vs Type I error cutoff α in Hypothesis Testing

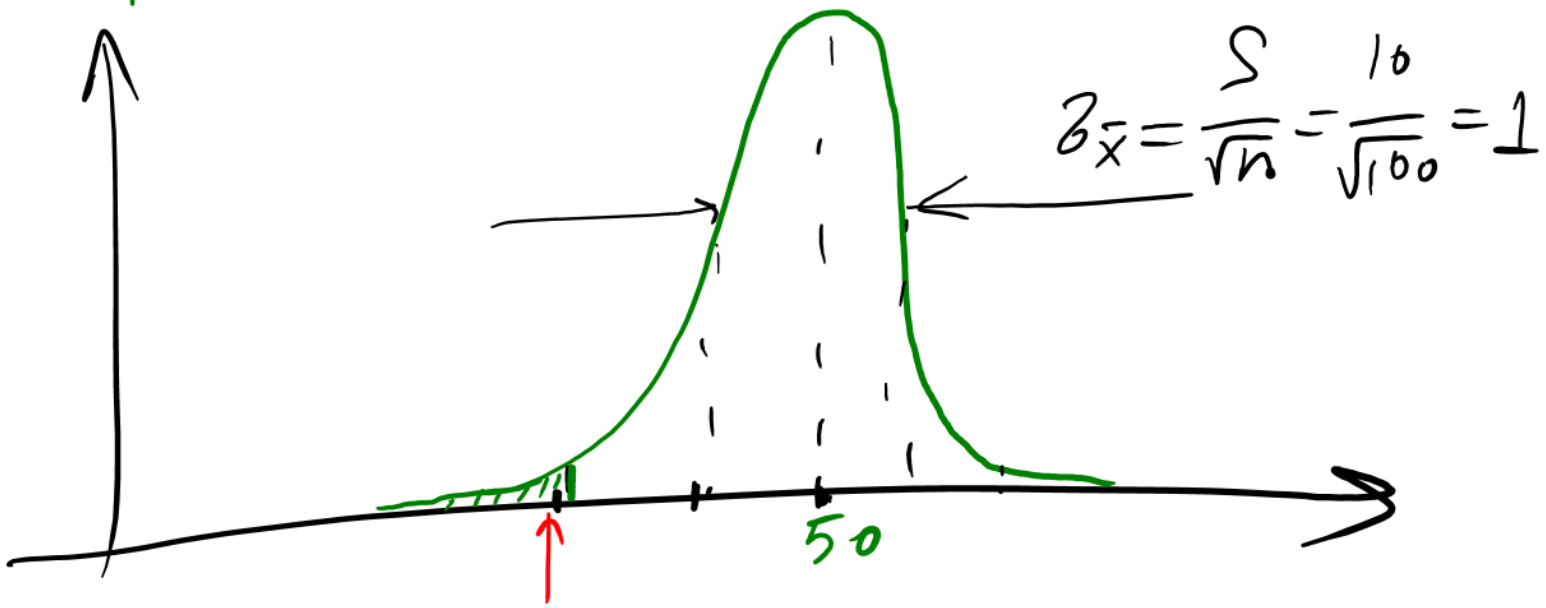
- Problem with using a predefined α: you don't know by how much you exceeded it
- Another approach is to calculate Prob(H_0 is true given the sample data) referred to as P-value. It the smallest α that would lead to rejection of null hypothesis
- You give your boss the P-value and let him/her decide if it is good enough
- Routinely with big datasets in genomics and systems biology P-values can be $10^{-\text{large number}\sim 10\text{-}100}$. This number is used to judge the quality of the hypothesis

$\mu_h = 50$

$H_0: \mu_c = \mu_h$

$n = 100, \quad \bar{X} = 48, \quad S = 10$

$H_1: \mu_c < \mu_h$



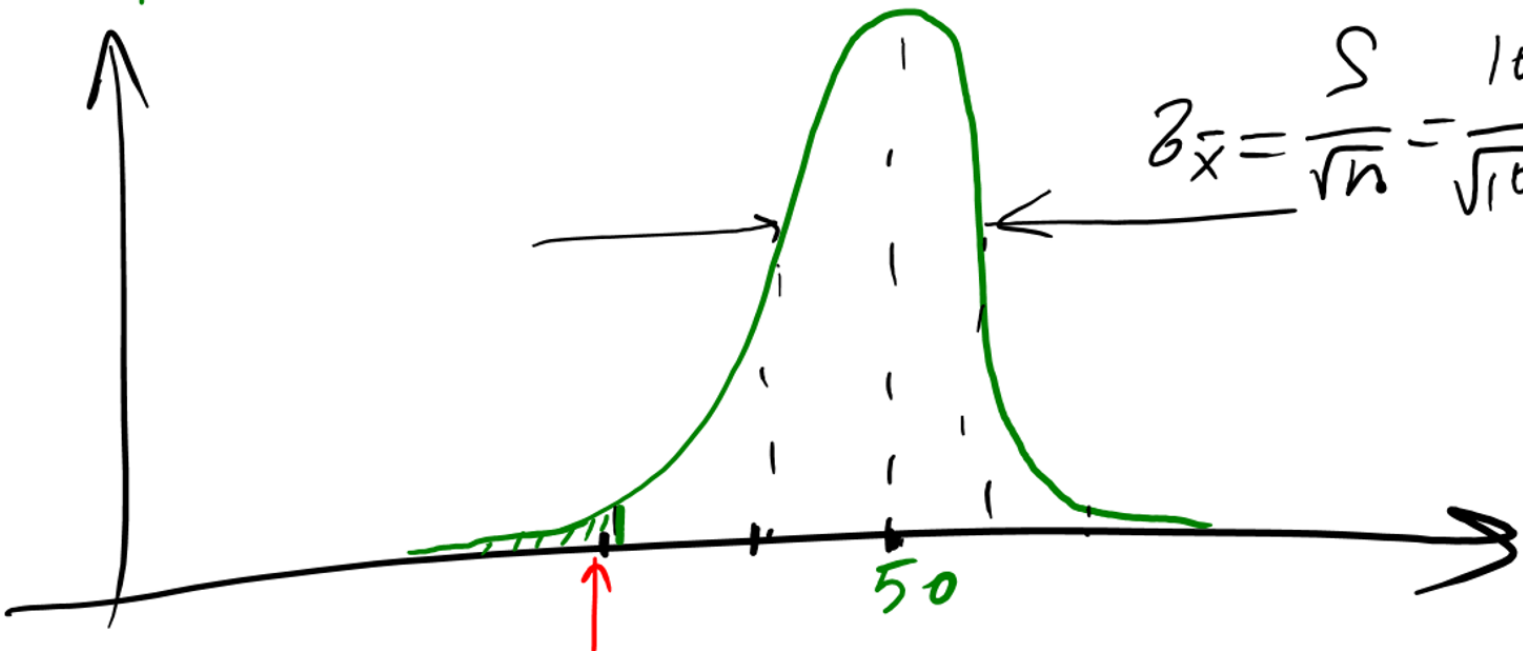
$\mu_h = 50$

$H_0: \mu_c = \mu_h$

$n = 100, \quad \bar{X} = 48, \quad S = 10$

$H_1: \mu_c < \mu_h$

$$\sigma_{\bar{x}} = \frac{S}{\sqrt{n}} = \frac{10}{\sqrt{100}} = 1$$

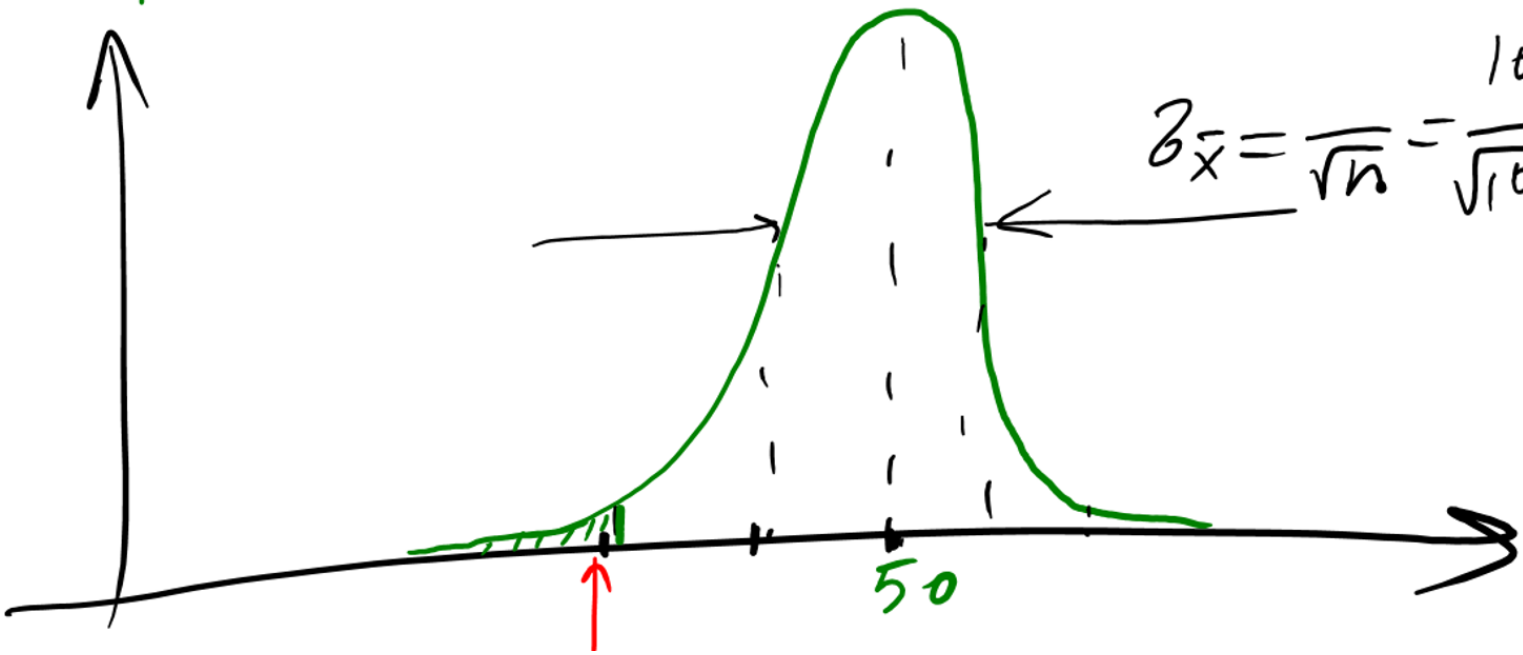


$\mu_h = 50$

$H_0: \mu_c = \mu_h$

$n = 100, \quad \bar{X} = 48, \quad \sigma = 10$

$H_1: \mu_c < \mu_h$



$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{100}} = 1$$

Generalizations

- What if H_1 is a two-sided hypothesis?
- A: P-value is $2(1-\Phi(|Z|))$, where $Z=(\bar{X}-\mu_0)/[S/\sqrt{n}]$

  Compare it to: For one sized $\mu_1>\mu_0$ it is $1-\Phi(Z)$

  For one sized $\mu_1<\mu_0$ it is $\Phi(Z)$
- If α is given, use μ_0 +/- $z_{\alpha/2}$*S as thresholds

to reject the null hypothesis

- What if the sample size n is small (say n<10):
- A: Use t-distribution with n-1 degrees of freedom for 2-sided *P-value=2(1-CDF_Tdist*(|*T*|)) where $T=(\bar{X}-\mu_0)/[S/\sqrt{n}]$.
- For a given α use μ_0 +/- $t_{\alpha/2,n-1}$ *T* to reject the null hypothesis

Type II Error and Choice of Sample Size

Assume you know the minimum $\delta = |\mu_1 - \mu_0|$ that you care about.
What is the minimal sample you should use to separate H0 and H1 hypotheses if your tolerance to type I and type II errors is α and β ?

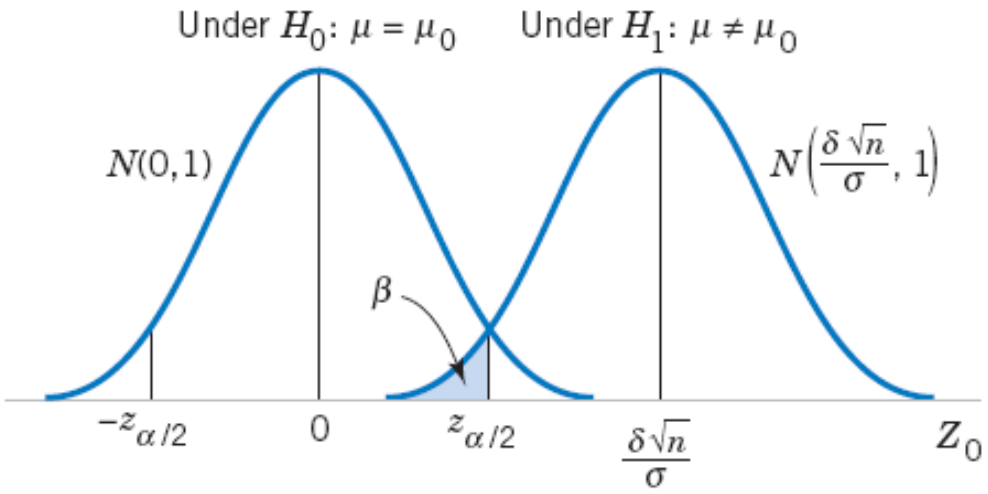


Figure 9-9 The distribution of Z_0 under H_0 and H_1.

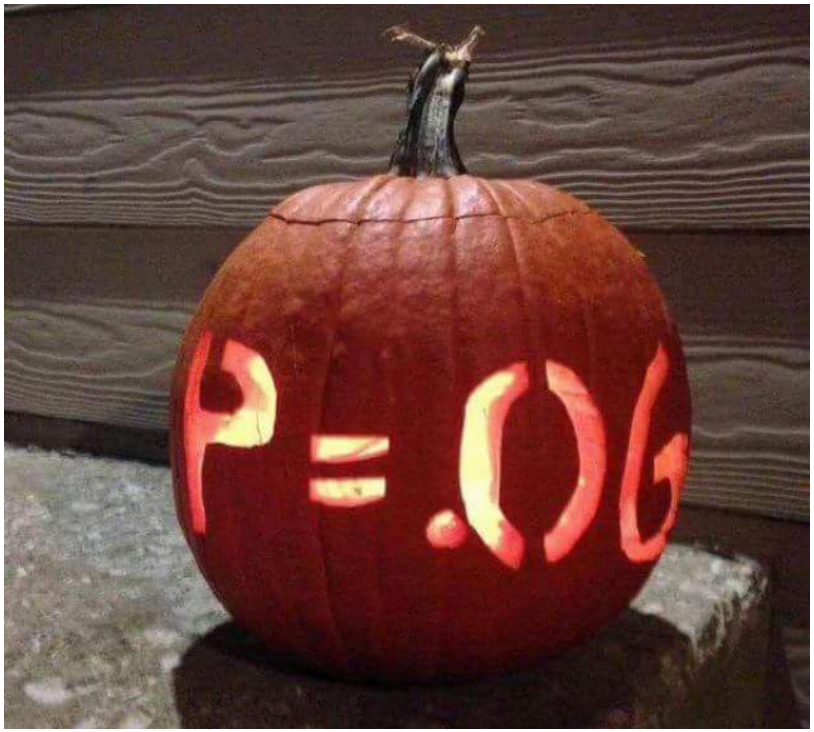
$$\frac{\delta\sqrt{n}}{\sigma} = z_{\alpha/2} + z_{\beta}$$

$$n \simeq \frac{(z_{\alpha/2} + z_{\beta})^2 \sigma^2}{\delta^2} \qquad \text{where} \qquad \delta = \mu - \mu_0 \qquad (9\text{-}22)$$

Standard notation to indicate P-value with *, **, ***

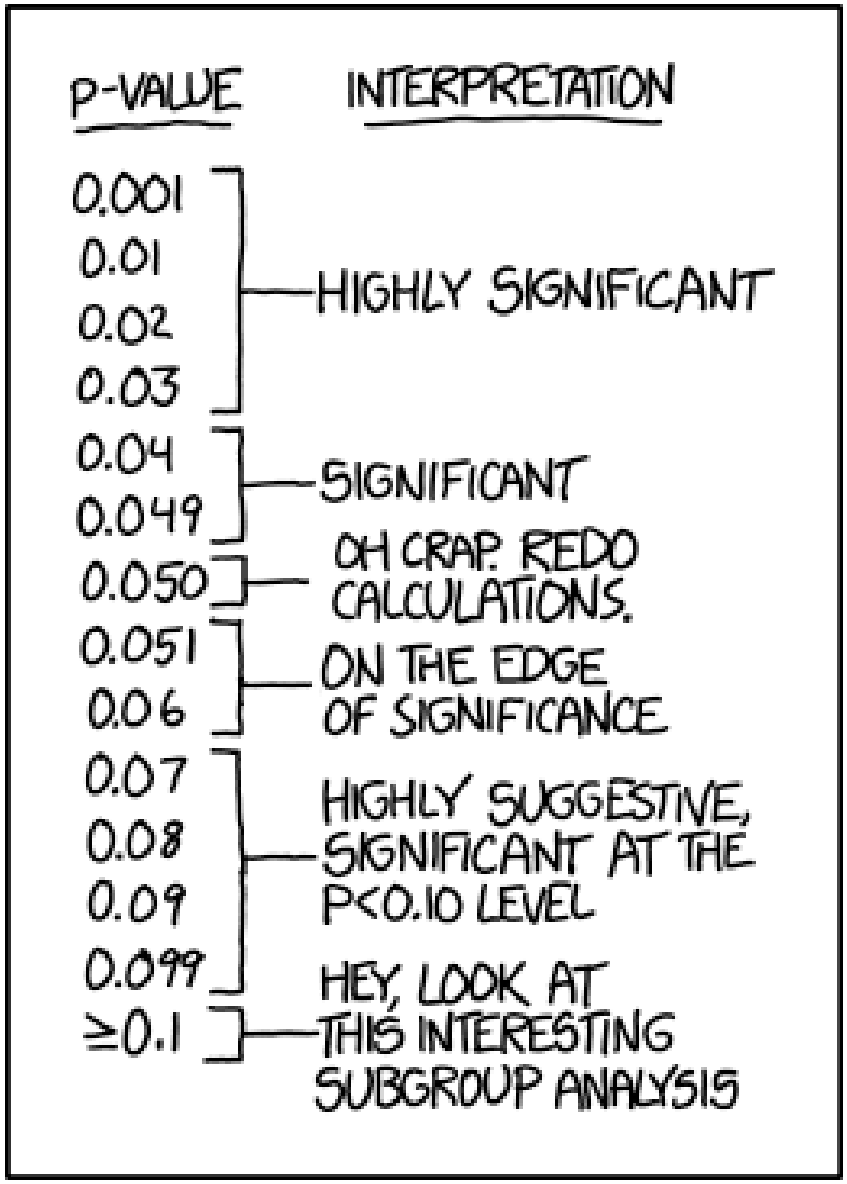
Table 11.1: A commonly adopted convention for reporting p values: in many places it is conventional to report one of four different things (e.g., $p < .05$) as shown below. I've included the "significance stars" notation (i.e., a * indicates $p < .05$) because you sometimes see this notation produced by statistical software. It's also worth noting that some people will write *n.s.* (not significant) rather than $p > .05$.

Usual notation	Signif. stars	English translation	The null is...
$p > .05$		The test wasn't significant	Retained
$p < .05$	*	The test was significant at $\alpha = .05$ but not at $\alpha = .01$ or $\alpha = .001$.	Rejected
$p < .01$	**	The test was significant at $\alpha = .05$ and $\alpha = .01$ but not at $\alpha = .001$.	Rejected
$p < .001$	***	The test was significant at all levels	Rejected



Happy Halloween! (belated)

Credit: Trust me, I'm a "Biologist" Facebook community



Credit: XKCD comics

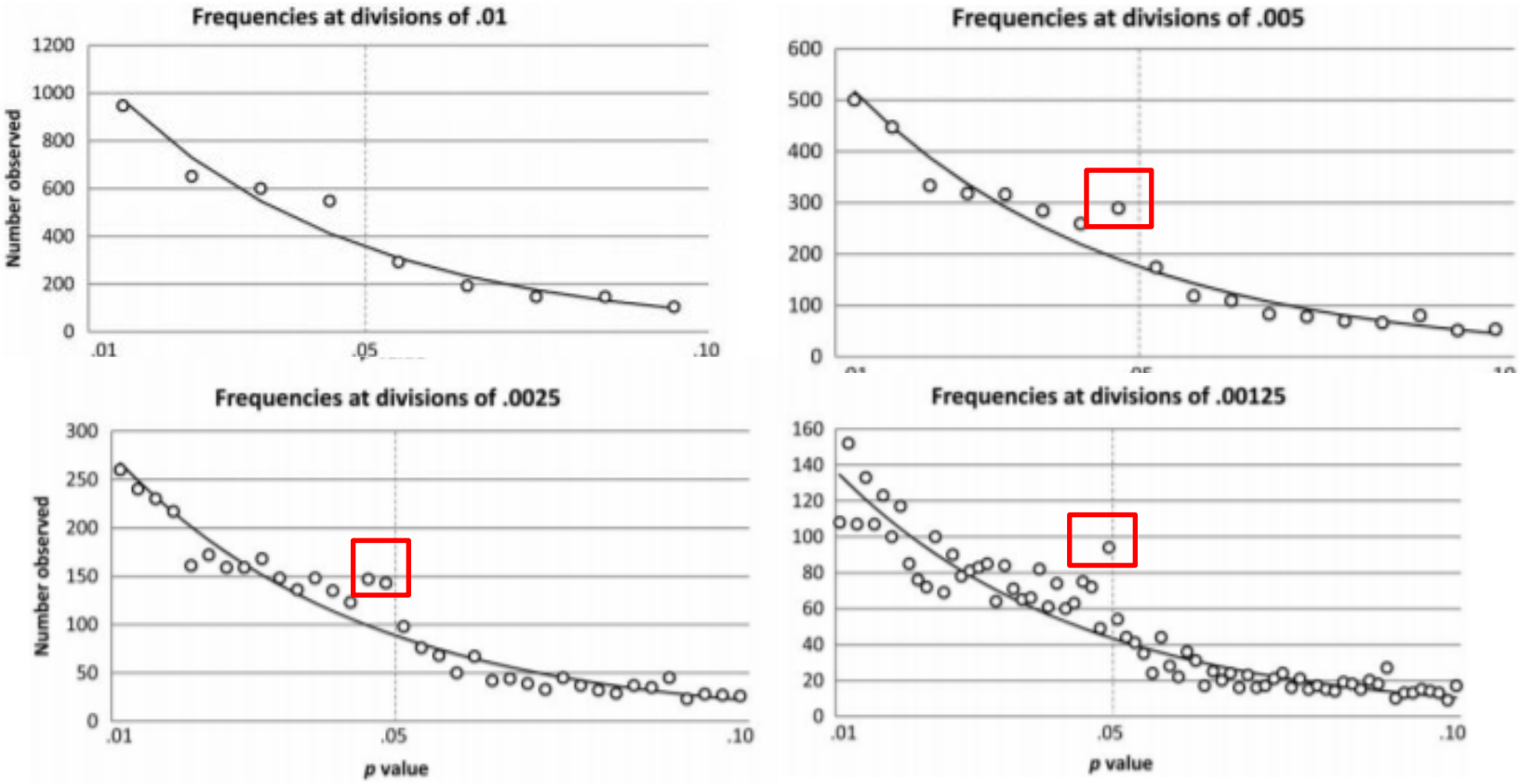
THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2012, 65 (11), 2271–2279

Psychology Press
Taylor & Francis Group

A peculiar prevalence of *p* values just below .05

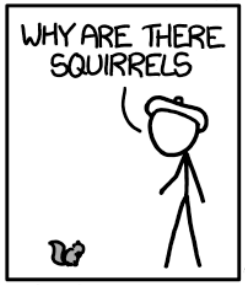
E. J. Masicampo[1], and Daniel R. Lalande[2]

[1]Department of Psychology, Wake Forest University, Winston-Salem, NC, USA
[2]Department of Health Sciences, Université du Québec à Chicoutimi, Chicoutimi, QC, Canada



Credit: XKCD comics

QUESTIONS
FOUND IN GOOGLE AUTOCOMPLETE

WHY DO WHALES JUMP
WHY ARE WITCHES GREEN
WHY ARE THERE MIRRORS ABOVE BEDS
WHY DO I SAY UH
WHY IS SEA SALT BETTER
WHY ARE THERE TREES IN THE MIDDLE OF FIELDS
WHY IS THERE NOT A POKEMON MMO
WHY IS THERE LAUGHING IN TV SHOWS
WHY ARE THERE DOORS ON THE FREEWAY
WHY ARE THERE SO MANY SVCHOST.EXE RUNNING
WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA
WHY ARE THERE SCARY SOUNDS IN MINECRAFT
WHY IS THERE KICKING IN MY STOMACH
WHY ARE THERE TWO SLASHES AFTER HTTP
WHY ARE THERE CELEBRITIES
WHY DO SNAKES EXIST
WHY DO OYSTERS HAVE PEARLS
WHY ARE DUCKS CALLED DUCKS
WHY DO THEY CALL IT THE CLAP
WHY ARE KYLE AND CARTMAN FRIENDS
WHY IS THERE AN ARROW ON AANG'S HEAD
WHY ARE TEXT MESSAGES BLUE
WHY ARE THERE MUSTACHES ON CLOTHES
WHY ARE THERE MUSTACHES ON CARS
WHY ARE THERE MUSTACHES EVERYWHERE
WHY ARE THERE SO MANY BIRDS IN OHIO
WHY IS THERE SO MUCH RAIN IN OHIO
WHY IS OHIO WEATHER SO WEIRD
WHY ARE THERE MALE AND FEMALE BIKES
WHY ARE THERE BRIDESMAIDS
WHY DO DYING PEOPLE REACH UP
WHY AREN'T THERE VARICOSE ARTERIES
WHY ARE OLD KLINGONS DIFFERENT
WHY ARE THERE SQUIRRELS

WHY IS PROGRAMMING SO HARD
WHY IS THERE A 0 OHM RESISTOR
WHY DO AMERICANS HATE SOCCER
WHY DO RHYMES SOUND GOOD
WHY DO TREES DIE
WHY IS THERE NO SOUND ON CNN
WHY AREN'T POKEMON REAL
WHY AREN'T BULLETS SHARP
WHY DO DREAMS SEEM SO REAL
WHY ARE THERE TINY SPIDERS IN MY HOUSE
WHY DO SPIDERS COME INSIDE
WHY ARE THERE HUGE SPIDERS IN MY HOUSE
WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE
WHY ARE THERE SPIDERS IN MY ROOM
WHY ARE THERE SO MANY SPIDERS IN MY ROOM
WHY DO SPIDER BITES ITCH
WHY IS DYING SO SCARY
WHY IS THERE NO GPS IN LAPTOPS
WHY DO KNEES CLICK
WHY AREN'T THERE E GRADES
WHY IS ISOLATION BAD
WHY DO BOYS LIKE ME
WHY DON'T BOYS LIKE ME
WHY IS THERE ALWAYS A JAVA UPDATE
WHY ARE THERE RED DOTS ON MY THIGHS
WHY IS LYING GOOD
WHY IS SEX SO IMPORTANT

WHY ARE THERE SLAVES IN THE BIBLE
WHY DO TWINS HAVE DIFFERENT FINGERPRINTS
WHY ARE AMERICANS AFRAID OF DRAGONS
WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT
WHY AREN'T MY ARMS GROWING

WHY ARE THERE SWARMS OF GNATS
WHY IS THERE PHLEGM
WHY ARE THERE SO MANY CROWS IN ROCHESTER,
WHY AREN'T ECONOMISTS RICH
WHY DO AMERICANS CALL IT SOCCER
WHY ARE MY EARS RINGING
WHY ARE THERE SO MANY AVENGERS
WHY ARE THE AVENGERS FIGHTING THE X MEN
WHY IS WOLVERINE NOT IN THE AVENGERS
WHY IS PSYCHIC WEAK TO BUG
WHY DO CHILDREN GET CANCER
WHY IS POSEIDON ANGRY WITH ODYSSEUS
WHY IS THERE ICE IN SPACE
WHY ARE THERE ANTS IN MY LAPTOP
WHY IS EARTH TILTED
WHY IS SPACE BLACK
WHY IS OUTER SPACE SO COLD
WHY ARE THERE PYRAMIDS ON THE MOON
WHY IS NASA SHUTTING DOWN
WHY ARE THERE GHOSTS

WHY IS THERE AN OWL IN MY BACKYARD
WHY IS THERE AN OWL OUTSIDE MY WINDOW
WHY IS THERE AN OWL ON THE DOLLAR BILL
WHY DO OWLS ATTACK PEOPLE
WHY ARE AK 47s SO EXPENSIVE
WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE
WHY ARE THERE GODS
WHY ARE THERE TWO SPOCKS
WHY ARE MY BOOBS ITCHY
WHY ARE CIGARETTES LEGAL
WHY ARE THERE DUCKS IN MY POOL
WHY IS JESUS WHITE
WHY IS THERE LIQUID IN MY EAR
WHY DO Q TIPS FEEL GOOD
WHY DO GOOD PEOPLE DIE
WHY IS MT VESUVIUS THERE
WHY DO THEY SAY T MINUS
WHY ARE THERE OBELISKS
WHY ARE WRESTLERS ALWAYS WET
WHY ARE OCEANS BECOMING MORE ACIDIC
WHY IS ARWEN DYING
WHY AREN'T MY QUAIL LAYING EGGS
WHY AREN'T MY QUAIL EGGS HATCHING
WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA
WHY AREN'T THERE GUNS IN HARRY POTTER

WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

WHY DO IGUANAS DIE
WHY AREN'T THERE DINOSAUR GHOSTS
WHY IS THERE HELL IF GOD FORGIVES
WHY IS GPS FREE
WHY ARE THERE FEMALE MR MIMES
WHY IS THERE LAVA
WHY IS YKK ON ALL ZIPPERS
WHY ARE THERE WEEKS
WHY DO I FEEL DIZZY
WHY IS THERE NO KING IN ENGLAND
WHY ARE DOGS AFRAID OF FIREWORKS