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?

For blue M&Ms $p = 0.24$

$$1.96\sqrt{\frac{0.24(1-0.24)}{n}} < 0.04$$

$$n > \left(\frac{1.96}{0.04}\right)^2 0.24 \times (1-0.24) = 438 \text{ M\&Ms or}$$

~2 × 7oz bags with 210 candies each

For red M&Ms $p = 0.13$

$$n > \left(\frac{1.96}{0.04}\right)^2 \times 0.13 \times (1-0.13) \simeq 271 \text{ M\&Ms or}$$

~1 × 7oz bag

Example 10-7 **Chocolate and Cardiovascular Health** An article in *Nature* (2003, Vol. 48, p. 1013) described an

chocolate c
In the expe
late per day
consisted o
average bo

Is there ev
plasma ant

Plasma antioxidants from chocolate

Dark chocolate may offer its consumers health benefits the milk variety cannot match.

There is some speculation that dietary flavonoids from chocolate, in particular (−)epicatechin, may promote cardiovascular health as a result of direct antioxidant effects or through antithrombotic mechanisms[1–3]. Here we show that consumption of plain, dark chocolate (Fig. 1) results in an increase in both the total antioxidant capacity and the (−)epicatechin content of blood plasma, but that these effects are markedly reduced when the chocolate is consumed with milk or if milk is incorporated as milk chocolate. Our findings indicate that milk may interfere with the absorption of antioxidants from chocolate *in vivo* and may therefore negate the potential health benefits that can be derived from eating moderate amounts of dark chocolate.

To determine the antioxidant content of different chocolate varieties, we took dark chocolate and milk chocolate prepared from the same batch of cocoa beans and defatted them twice with *n*-hexane before extracting them with a mixture of water, acetone and acetic acid (70.0:29.8:0.2 by volume). We measured their *in vitro* total antioxidant capacities using the ferric-reducing antioxidant potential (FRAP) assay[4]; FRAP reduced iron per 100 g for dark and milk chocolate, respectively. Volunteers must therefore consume twice as much milk chocolate as dark chocolate to receive a similar intake of antioxidants.

We recruited 12 healthy volunteers (7 women and 5 men with an average age of 32.2 ± 1.0 years (range, 25–35 years). Subjects were non-smokers, had normal blood lipid levels, were taking no drugs or vitamin supplements, and had an average weight of 65.8 ± 3.1 kg (range, 46.0–86.0 kg) and body-mass index of 21.9 ± 0.4 kg m^{-2} (range, 18.6–23.6 kg m^{-2}). On different days, following a crossover experimental design, subjects consumed 100 g dark chocolate, 100 g dark chocolate with 200 ml full-fat milk, or 200 g milk chocolate (containing the equivalent of up to 40 ml milk).

One hour after subjects had ingested the chocolate, or chocolate and milk, we measured the total antioxidant capacity of their plasma by FRAP assay. Plasma antioxidant levels increased significantly after consumption of dark chocolate alone, from $100 \pm 3.5\%$ to $118.4 \pm 3.5\%$ (*t*-test, $P < 0.001$), returning to baseline values ($95.4 \pm 3.6\%$) after 4 h (Fig. 2a). There was

ION.COM/ALAMY

Mauro Serafini*, Rossana Bugianesi*, Giuseppe Maiani*, Silvia Valtuena*, Simone De Santis*, Alan Crozier†
**Antioxidant Research Laboratory, Unit of Human Nutrition, National Institute for Food and Nutrition Research, Via Ardeatina 546, 00178 Rome, Italy*
e-mail: serafini@inran.it
†*Plant Products and Human Nutrition Group, Graham Kerr Building, Division of Biochemistry and Molecular Biology, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK*

Figure 1 Stack of benefits? Unlike its milky counterpart, dark chocolate may provide more than just a treat for the tastebuds.

could be due to the formation of secondary bonds between chocolate flavonoids and milk proteins[6,7], which would reduce the biological accessibility of the flavonoids and therefore the chocolate's potential antioxidant properties *in vivo*.

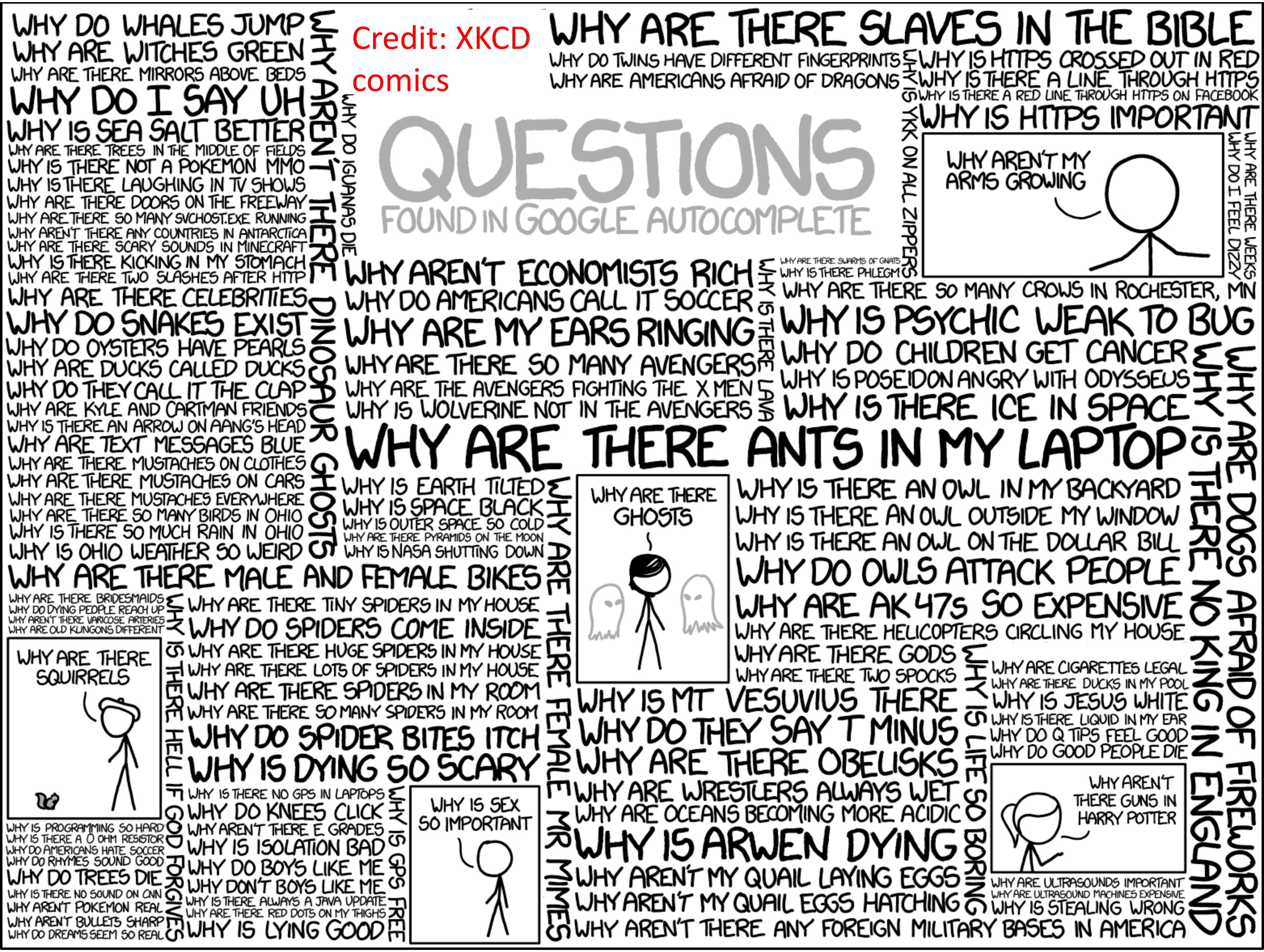
Our findings highlight the possibility

Sweet matlab exercise #1

- **Download dark_vs_milk_chocolate_analysis_template.m at the course website. Correct all ?? In the file**
- **dark=[118.8 122.6 115.6 113.6 119.5 115.9 115.8 115.1 116.9 115.4 115.6 107.9];**
- **milk=[102.1 105.8 99.6 102.7 98.8 100.9 102.8 98.7 94.7 97.8 99.7 98.6]**
- **Use Z-statistics to calculate P-value of the null hypothesis H_0 that milk = dark against H_1 that dark > milk. P_value_z=1-normcdf(Z)**
- **Repeat using T-statistics. # of degrees of freedom is dof=2*(n-1)**
  P_value_t=tcdf(T, dof)

Sweet matlab exercise #1

- **dark=[118.8 122.6 115.6 113.6 119.5 115.9 115.8 115.1 116.9 115.4 115.6 107.9];**
- **milk=[102.1 105.8 99.6 102.7 98.8 100.9 102.8 98.7 94.7 97.8 99.7 98.6]**
- **x_dark=mean(dark) % sample mean dark chocolate**
- **x_milk=mean(milk) % sample mean milk chocolate**
- **s_dark=std(dark) % sample std dark chocolate**
- **s_milk=std(milk) % sample std milk chocolate**
- **n=12 % sample size of both dark and milk**
- **std_xdiff=sqrt(s_dark.^2./2+s_milk.^2./n) % std diff x**
- **z_stat=(x_dark-x_milk)./std_xdiff % z-statistic**
- **P_value_z=erfc(z_stat./sqrt(2))./2 % P-value of null true**
- **% P_value_z=9.9629e-34**
- **dof=(n-1)+(n-1) % # of degrees of freedom**
- **P_value_t=tcdf(z_stat,dof,'upper') % P-value of null true**
- **%P_value_t= 1.8417e-11**



Credit: XKCD comics

Goodness of Fit hypothesis testing:

Pearson's chi-square test

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 accept or reject the null hypothesis that these probabilities are correct from a finite sample?

Pearson chi^2 Goodness of Fit Test

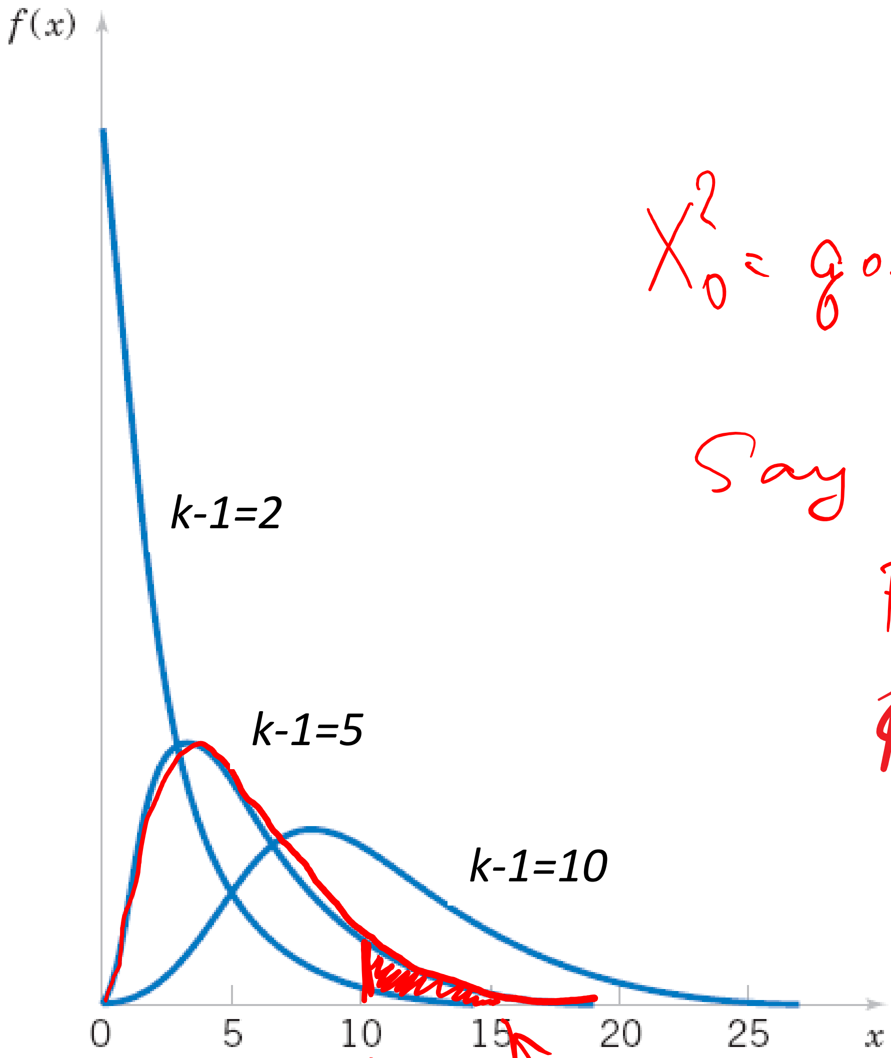
- Assume there is a sample of size n from a population with k classes (e.g. 6 M&M colors)
- Null hypothesis H_0: class i has frequency f_i in the population
- Alternative hypothesis H_1: some population frequencies are inconsistent with f_i
- Let O_i be the observed number of sample elements in the ith class and $E_i = n\, f_i$ be the expected number of sample elements in the ith class.
- Group any bin with E_i <3 with
  a) if numerical value of i is important, group it with its neighbor (k=i-1 or k=i+1) which has the smallest E_k until E_{group} >=3;
  b) If numerical value of i is irrelevant, group together all E_i <3 bins until E_{group} >=3
- The test statistic is

$$X_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \qquad (9\text{-}47)$$

P-value is calculated based on the chi-square distribution with k-1 degrees of freedom:

P-value = Prob(H_0 is correct) =1-CDF_chi-squared(X_0^2, k-1)

chi^2 Goodness of Fit Test is a *one-sided* hypothesis



$$X_0^2 = gof = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

Say $X_0^2 = 10$

For M&M

$k = 6 \rightarrow k-1 = 5$

X_0^2

p-Value that null hypothesis is correct

M&M group exercise

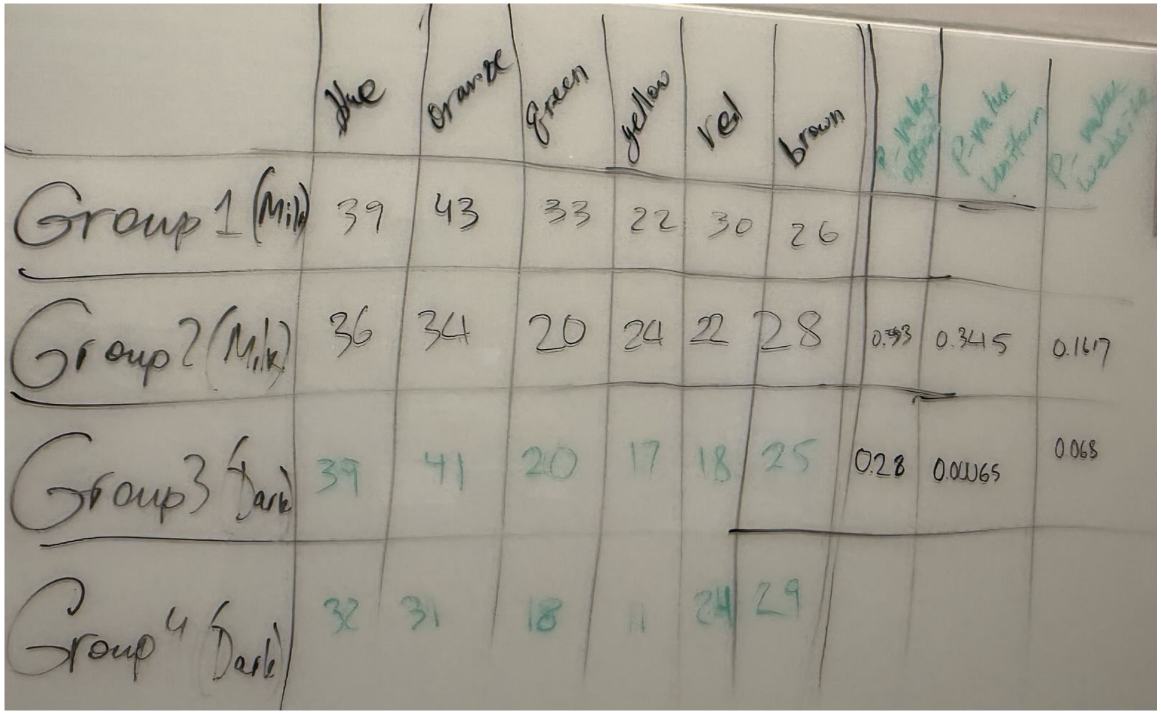
- DO NOT EAT CANDY BEFORE COUNTING IS FINISHED! THEN, **PLEASE, DO**.
- We will be testing three null hypotheses one after another:
  - M&M peanut official data: 23% blue, 23% orange, 15% green, 15% yellow, 12% red, 12% brown
  - Website (fan collected) data from http://joshmadison.com/2007/12/02/mms-color-distribution-analysis: 18.36% blue, 20.76% orange, 18.44% green, 14.08% yellow, 14.20% red, 14.16% brown
  - Uniform distribution: 1/6~16.67% of each candy color
- You will estimate P-values for each one of these null hypotheses
- Hints: O_i – is the observed # of candies of color i; calculate the expected # E_i=(# candies in your sample)*f_i

Use **1-chi2cdf(X0squared, 5)** for P-value

$$X_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

M&M matlab exercise

```
observed=mm_table(group,:); group % use when analyzing one group
f_mm=[0.24,0.2,0.16, 0.14, 0.13,0.13];
f_u=1./6.*ones(1,6);
f_website=[18,21,18,14,14,14,14];
f_website=f_website./sum(f_website);
%p_website=[0.1836, 0.2076, 0.1844, 0.1408, 0.1420, 0.1416]
%p_u=[0.1500, 0.2200, 0.2100, 0.1200, 0.1600, 0.1500];
n=sum(observed)
expected_u=n.*f_u;
expected_mm=n.*f_mm;
expected_website=n.*f_website;
gf_mm=0; gf_u=0; gf_website=0;
for m=1:6;
   gf_mm=gf_mm+(observed(m)...
      -expected_mm(m)).^2./expected_mm(m);
   gf_u=gf_u+(observed(m)-expected_u(m)).^2./expected_u(m);
   gf_website=gf_website+(observed(m)...
      -expected_website(m)).^2./expected_website(m);
end;
disp('goodness of fit of MM ='); disp(num2str(gf_mm));
disp('p-value of MM ='); disp(num2str(1-chi2cdf(gf_mm,5))); disp(' ');
disp('goodness of fit of website ='); disp(num2str(gf_website));
disp('p-value of MM ='); disp(num2str(1-chi2cdf(gf_website,5))); disp(' ');
disp('goodness of fit of uniform ='); disp(num2str(gf_u));
disp('p-value of uniform='); disp(num2str(1-chi2cdf(gf_u,5)));
```



Groups 1 and 2 – Milk chocolate
They accept all three hypotheses with uniform being weaker (smaller P-value but >0.05)

Groups 3 and 4 – Dark chocolate
They reject all three hypotheses except group 3 confirming the official hypothesis

Statistical tests of independence

How to **test the hypothesis** if multiple sample are drawn from the same population?

- Table: samples (Student groups) – rows, classes (M&M colors) – columns
- Test if color fractions are **independent** from group
- ***P(Group 1 and Color = green) = P(Group 1)*P(Color green)***
- Compute for all groups/colors 6*4=24 in our case

  $E_{green}(group\ 1) = n_{tot} * (group1/n_{tot}) * (green/n_{tot})$

- $\chi^2 = \sum_{groups\,\&\,colors}^{n_{tot}} \frac{(O_{color}(group) - E_{color}(group))^2}{E_{color}(group)}$
- # degrees of freedom=(colors-1)*(groups-1)

- M&M exercise Fall 2025
- Was the Milk M&M bag mixed well?
  Let's compare the data from groups 1 and 2
- Was the Dark M&M bag mixed well?
  Let's compare the data from groups 3 and 4

	Blue	Orange	Green	Yellow	Red	Brown
Group 1 (milk)	39	43	33	22	30	26
Group 2 (milk)	36	34	20	24	22	28
Group 3 (dark)	39	41	20	17	18	25
Group 4 (dark)	32	31	18	11	24	29

- Using $\chi^2 = \sum_{groups\,\&\,colors} \frac{(O_{color}\,(group) - E_{color}(group))^2}{E_{color}(group)}$

with # degrees of freedom (colors-1)*(groups-1)
Find P-value of null hypothesis H_0 that
<u>samples are independent</u> from each other

- For milk the P-value is 0.6356. I mixed this bag well!
- For dark the P-value is 0.5646. This bag was already mixed

Was the bag of milk M&Ms well mixed?

- `clear mm_table`
- `mm_table=mm_table_all(1:2,:);`
- `ngroups=2;`
- `ncolors=6;`
- `sumt=sum(sum(mm_table))`
- `sum_color=sum(mm_table, 1)`
- `sum_group=sum(mm_table, 2)`
- `mm_exp=kron(sum_group,sum_color)./sumt`
- `gof=sum(sum((mm_table-mm_exp).^2./mm_exp))`
- `P_value_gof=1-chi2cdf(gof, (ngroups-1)*(ncolors-1))`
- **`%gof = 3.4192`**
- **`P_value_gof = 0.6356`**
- The null model that samples are independent is **not rejected → The bag was well mixed!**

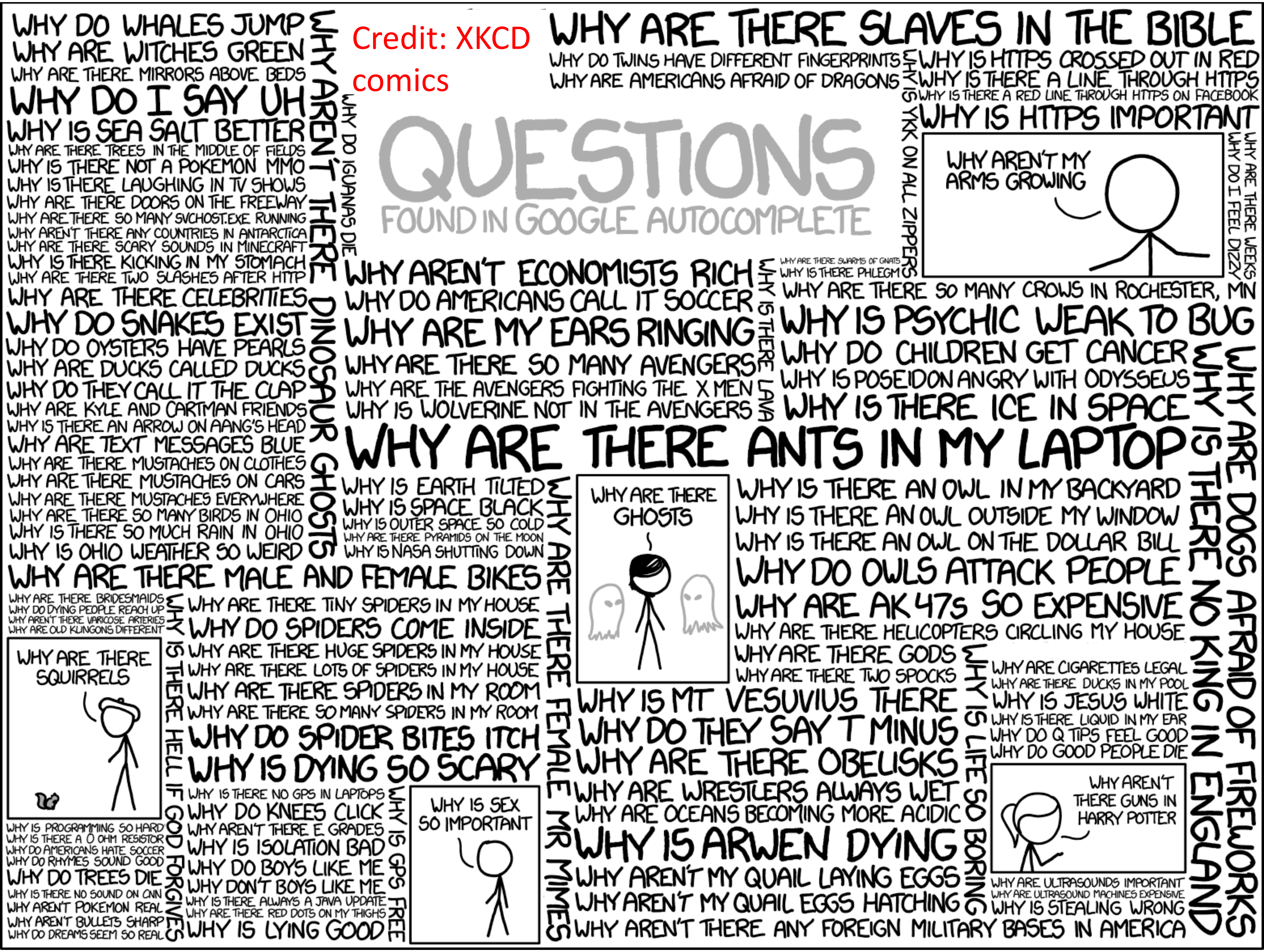
Batch effect

Does color composition vary between Milk and Dark chocoltae M&M

- Milk total: groups 1+2
  75 77 53 46 52 54
- Dark chocolate: groups 3+4
  71 72 38 28 42 54
- Test if they are significantly different from each other:
- Same statistical independence test:
  ngroups=2; ncolors=6;
- Results:

  Goodness of Fit = 4.1330

  P_value_gof = 0.5304
- Batch effect is not statistically significant. Milk and Dark M&Ms have the same color frequencies



Credit: XKCD comics

Goodness of fit with a PDF defined by m parameters

- As before: k classes (e.g. M&M colors)
- Use parameter estimators to find the best parameters for the fit
  - Method of moments
  - MLE: method of maximum likelihood
- Use chi-squared distribution with k-1-m degrees of freedom
- As before: if $E_i < 3$, group it together with another group and reduce k by 1

$$X_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \qquad (9\text{-}47)$$

9-7 Testing for Goodness of Fit

Example 9-12

EXAMPLE 9-12 Printed Circuit Board Defects

Poisson Distribution

The number of defects in printed circuit boards is hypothesized to follow a Poisson distribution. A random sample of $n = 60$ printed boards has been collected, and the following number of defects observed.

Number of Defects	Observed Frequency
0	32
1	15
2	9
3	4

9-7 Testing for Goodness of Fit

Example 9-12

The mean of the assumed Poisson distribution in this example is unknown and must be estimated from the sample data. The estimate of the mean number of defects per board is the sample average, that is, $(32 \cdot 0 + 15 \cdot 1 + 9 \cdot 2 + 4 \cdot 3)/60 = 0.75$. From the Poisson distribution with parameter 0.75, we may compute p_i, the theoretical, hypothesized probability associated with the ith class interval. Since each class interval corresponds to a particular number of defects, we may find the p_i as follows:

$$p_1 = P(X = 0) = \frac{e^{-0.75}(0.75)^0}{0!} = 0.472$$

$$p_2 = P(X = 1) = \frac{e^{-0.75}(0.75)^1}{1!} = 0.354$$

$$p_3 = P(X = 2) = \frac{e^{-0.75}(0.75)^2}{2!} = 0.133$$

$$p_4 = P(X \geq 3) = 1 - (p_1 + p_2 + p_3) = 0.041$$

9-7 Testing for Goodness of Fit

Example 9-12

The expected frequencies are computed by multiplying the sample size $n = 60$ times the probabilities p_i. That is, $E_i = np_i$. The expected frequencies follow:

Number of Defects	Probability	Expected Frequency
0	0.472	28.32
1	0.354	21.24
2	0.133	7.98
3 (or more)	0.041	2.46

9-7 Testing for Goodness of Fit

Example 9-12

Since the expected frequency in the last cell is less than 3, we combine the last two cells:

Number of Defects	Observed Frequency	Expected Frequency
0	32	28.32
1	15	21.24
2 (or more)	13	10.44

The chi-square test statistic in Equation 9-47 will have $k - p - 1 = 3 - 1 - 1 = 1$ degree of freedom, because the mean of the Poisson distribution was estimated from the data.

9-7 Testing for Goodness of Fit

Example 9-12

The seven-step hypothesis-testing procedure may now be applied, using $\alpha = 0.05$, as follows:

1. **Parameter of interest:** The variable of interest is the form of the distribution of defects in printed circuit boards.
2. **Null hypothesis:** H_0: The form of the distribution of defects is Poisson.
3. **Alternative hypothesis:** H_1: The form of the distribution of defects is not Poisson.
4. **Test statistic:** The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(o_i - E_i)^2}{E_i}$$

9-7 Testing for Goodness of Fit

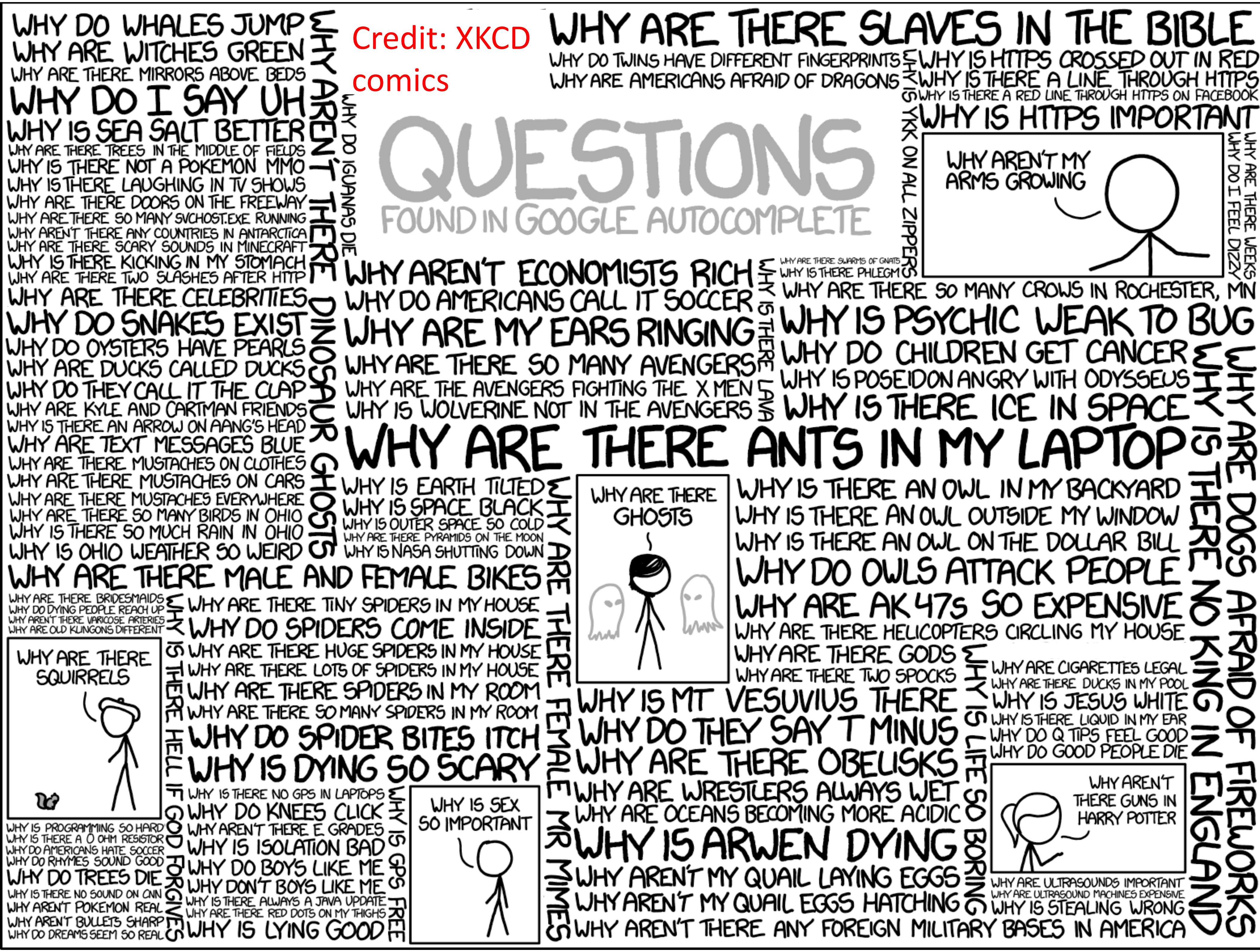
Example 9-12

5. **Reject H_0 if:** Reject H_0 if the P-value is less than 0.05.

6. **Computations:**

$$\chi_0^2 = \frac{(32 - 28.32)^2}{28.32} + \frac{(15 - 21.24)^2}{21.24} + \frac{(13 - 10.44)^2}{10.44} = 2.94$$

7. **Conclusions:** We find from Appendix Table III that $\chi^2_{0.10,1} = 2.71$ and $\chi^2_{0.05,1} = 3.84$. Because $\chi_0^2 = 2.94$ lies between these values, we conclude that the P-value is between 0.05 and 0.10. Therefore, since the P-value exceeds 0.05 we are unable to reject the null hypothesis that the distribution of defects in printed circuit boards is Poisson. The exact P-value computed from Minitab is 0.0864.



Credit: XKCD comics