Hypothesis testing: two samples

10-2: Inference for a Difference in Means of Two Normal Distributions, Variances Known

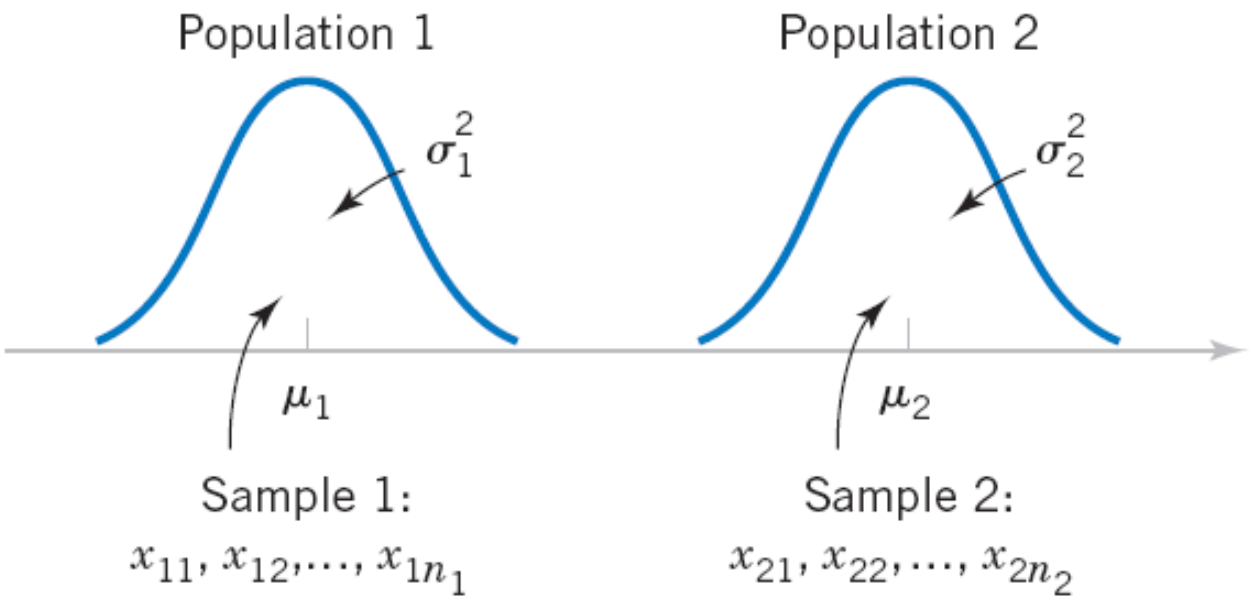


Figure 10-1 Two independent populations.

Figure 10-1 Two independent populations.

10-2: Inference for a Difference in Means of Two Normal Distributions, Variances Known

Assumptions

1. $X_{11}, X_{12}, \ldots, X_{1n_1}$ is a random sample from population 1.
2. $X_{21}, X_{22}, \ldots, X_{2n_2}$ is a random sample from population 2.
3. The two populations represented by X_1 and X_2 are independent.
4. Both populations are normal.

$$E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_1 - \mu_2$$

$$V(\bar{X}_1 - \bar{X}_2) = V(\bar{X}_1) + V(\bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

10-2: Inference for a Difference in Means of Two Normal Distributions, Variances Known

The quantity

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad (10\text{-}1)$$

has a $N(0, 1)$ distribution.

10-2: Inference for a Difference in Means of Two Normal Distributions, Variances Known

10-2.1 Hypothesis Tests for a Difference in Means, Variances Known

usually $\Delta_0 = 0$

Null hypothesis: $H_0: \mu_1 - \mu_2 = \Delta_0$

Test statistic:
$$Z_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad (10\text{-}2)$$

Alternative Hypotheses	*P*-Value	Rejection Criterion For for Fixed-Level Tests
$H_1: \mu_1 - \mu_2 \neq \Delta_0$	Probability above $\lvert z_0 \rvert$ and probability below $-\lvert z_0 \rvert$, $P = 2[1 - \Phi(\lvert z_0 \rvert)]$	$z_0 > z_{\alpha/2}$ or $z_0 < -z_{\alpha/2}$
$H_1: \mu_1 - \mu_2 > \Delta_0$	Probability above z_0, $P = 1 - \Phi(z_0)$	$z_0 > z_\alpha$
$H_1: \mu_1 - \mu_2 < \Delta_0$	Probability below z_0, $P = \Phi(z_0)$	$z_0 < -z_\alpha$

10-2.1 Hypotheses Tests on the Difference in Means, Variances Unknown

Case 2: $\sigma_1^2 \neq \sigma_2^2$

If H_0: $\mu_1 - \mu_2 = \Delta_0$ is true, the statistic

$$T_0^* = \frac{\overline{X}_1 - \overline{X}_2 - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \qquad (10\text{-}15)$$

is distributed as t-distribution with degrees of freedom given by

$$v = n_1 + n_2 - 2,$$

or more generally

Manhattan plot for Genome-Wide Association Study (GWAS)

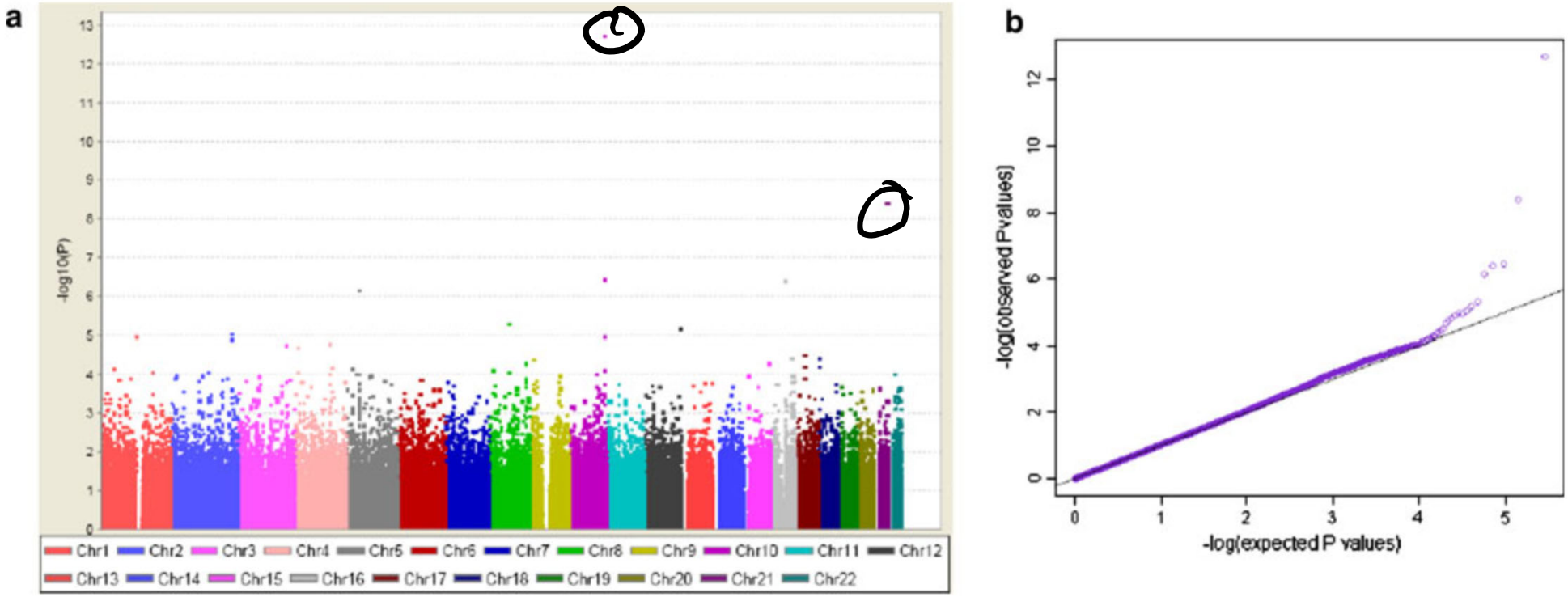


Fig. 1 Genome-wide association results comparing 2,702 cases and 5,726 controls: **a** Manhattan and **b** quantile–quantile plots of −log10 tranformed *P* values of 285,984 SNPS genotyped

Li J,, et al. A combined analysis of genome-wide association studies in breast cancer. Breast Cancer Res Treat. 2011;126: 717–727

Multiple null hypotheses: Bonferroni correction

- What if you have **m** independent null hypotheses? Say you have **m=25,000 genes** in a genome?
- What is the probability that at least one of the null-hypotheses will be shown to be false at significance threshold α_1?
- Answer: Family-Wise Error Rate or $FWER=1-(1-\alpha_1)^m \approx m\alpha_1$
- If m=20 and $\alpha_1=0.05$, $FWER= 0.6415$
- If you want to get $FWER< \alpha$, use $\alpha_1= \alpha/m$

Carlo Emilio Bonferroni
(1892 –1960)
Italian mathematician who worked on probability theory.



Example 10-7 Chocolate and Cardiovascular Health An article in *Nature* (2003, Vol. 424, p. 1013) described an

chocolate
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 ± 3.5% to 118.4 ± 3.5% (*t*-test, $P < 0.001$), returning to baseline values (95.4 ± 3.6%) after 4 h (Fig. 2a). There was



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

Vol. 424

© 2003 Nature Publishing Group

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=2*[1-normcdf(|Z|)]**
- **Repeat using T-statistics. # of degrees of freedom is dof=2*(n-1)**
  P_value_t=2*tcdf(|T|, dof)

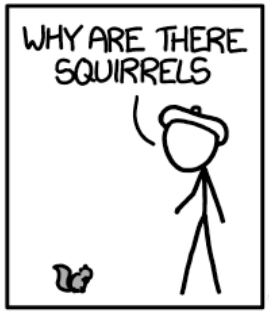
Sweet matlab exercise #1

- 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

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 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



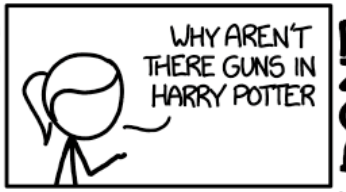
WHY ARE THERE SLAVES IN THE BIBLE
WHY DO TWINS HAVE DIFFERENT FINGERPRINTS
WHY ARE AMERICANS AFRAID OF DRAGONS
WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT



WHY ARE THERE SWARMS OF GNATS
WHY IS THERE PHLEGM
WHY ARE THERE SO MANY CROWS IN ROCHESTER, MN
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 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 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 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 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 LAVA
WHY ARE THERE FEMALE MR MIMES
WHY IS GPS FREE
WHY IS LIFE SO BORING
WHY IS THERE NO KING IN ENGLAND
WHY ARE DOGS AFRAID OF FIREWORKS
WHY ARE THERE WEEKS
WHY DO I FEEL DIZZY
WHY IS YKK ON ALL ZIPPERS
WHY IS THERE HELL IF GOD FORGIVES

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?



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 $p = 0.24$

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

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

$\sim$ 2 x 7oz bags with 210 candies each

For red M&Ms $p = 0.13$

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

$\sim$ 1 x 7oz bag

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

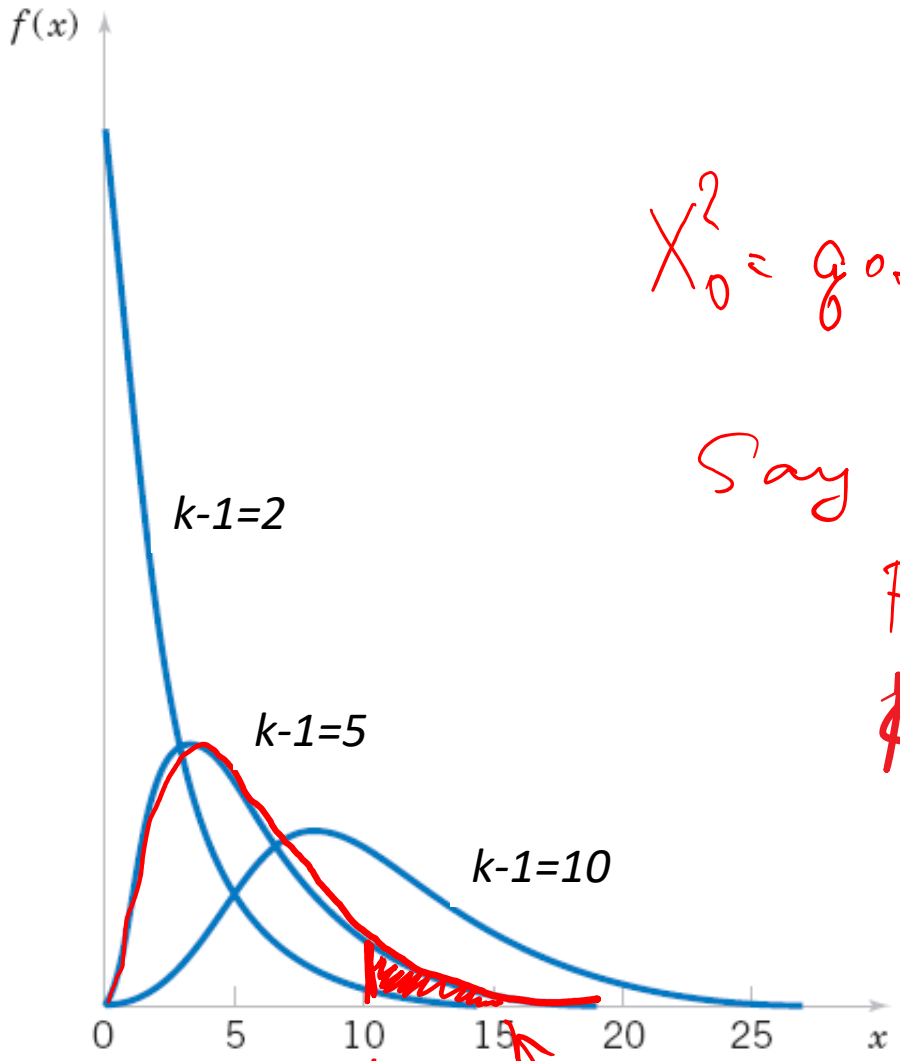
- Assume there is a sample of size n from a population with k classes (e.g. 6 M&M colors)
- Null hypothesis H_0: class i has frequency f_i in the population
- Alternative hypothesis H_1: some population frequencies are inconsistent with f_i
- Let O_i be the observed number of sample elements in the ith 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



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

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

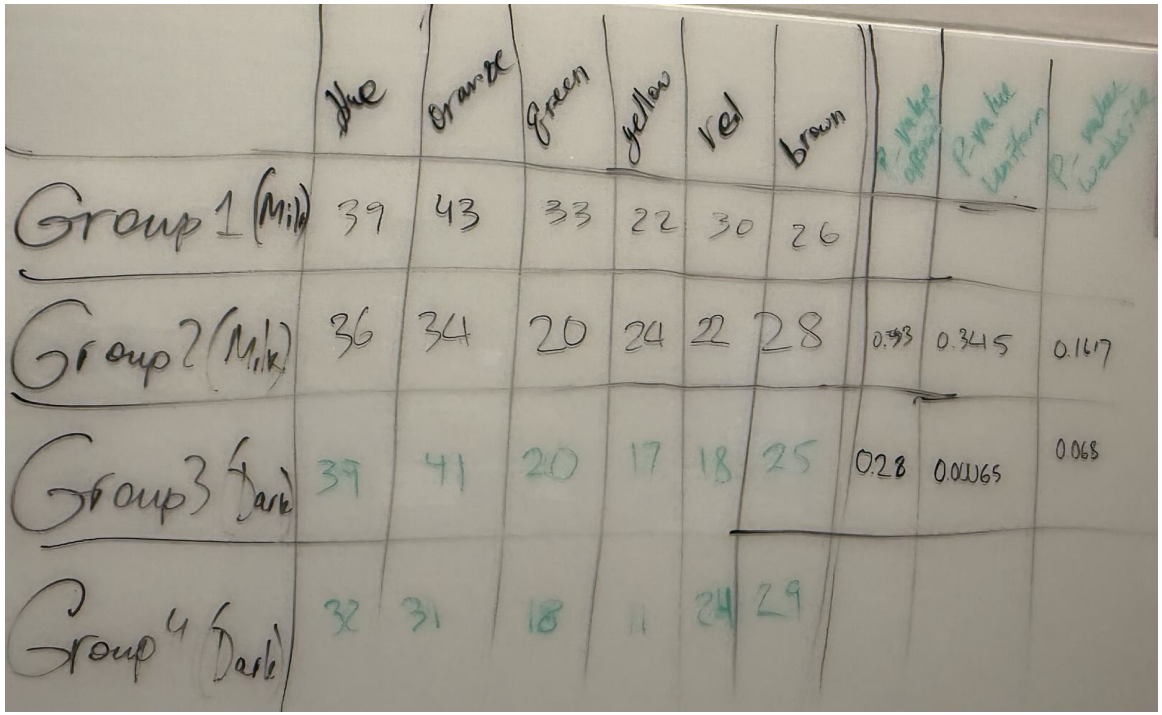
- DO NOT EAT CANDY BEFORE COUNTING IS FINISHED! THEN, **PLEASE, DO**.
- We will be testing three null hypotheses one after another:
  - M&M peanut official data: 23% blue, 23% orange, 15% green, 15% yellow, 12% red, 12% brown
  - Website (fan collected) data from http://joshmadison.com/2007/12/02/mms-color-distribution-analysis: 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