

SOLUTIONS

Alex Barnett
3/13/06

Math 50: Final


180 minutes, 140 points. No algebra-capable calculators. Try to use your calculator only at the end of your calculation, and show working/reasoning. Please *do* look up z , t , χ^2 values for full credit. Attempt all questions, but in the order most comfortable for you, and heed the number of points available. Later question parts may be independent of earlier ones, so skip over one you can't do.

1. [13 points] It was found that 35 out of 300 'famous' people have the star sign Sagittarius. not that Sagittarian is famous!
- 8 (a) Test the null hypothesis that Sagittariuses are no more likely to become famous than any other star sign (assume each sign is exactly 1/12 of the year), against the hypothesis that they are more prone to fame and fortune. (at 95% confidence level)

the binomial rand. var. X

$$H_0: p_0 = 1/12 \quad \text{vs.} \quad H_1: p > 1/12$$

under H_0 , X has binomial pdf with $n=300$, $p_0=1/12$, which we approximate by a normal with $\mu = np_0 = 25$
 $\sigma = \sqrt{np_0(1-p_0)} = 4.79$

z -value of observed data is $\frac{35 - \mu}{\sigma} = 2.09$ 

However, since k discrete (integer), continuity correction gives

$$z = \frac{34.5 - \mu}{\sigma} = 1.98 \quad \text{Either way, 1-sided } H_1 \text{ is accepted since } z > 1.64$$

- 5 (b) Give this same data, what would a Bayesian, with no prior bias towards any particular value of the proportion p of famous Sagittariuses, conclude about p ? (That is, what is the normalized posterior on p) $\neq z^{-1}(0.95)$
- or, strictly, I meant, Sagittarian famous-people

Uninformative prior means $g(p|k) = C \cdot p_x(k; p)$
posterior likelihood

$$= C \cdot p^{35} (1-p)^{300-35}$$

which is a beta pdf of parameters $r = 36$, $s = 266$

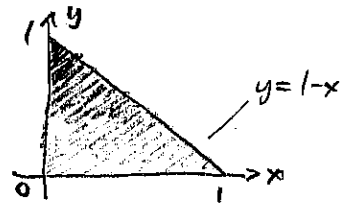
$$\Rightarrow g(p|k) = \frac{\Gamma(302)}{\Gamma(36)\Gamma(266)} p^{35} (1-p)^{265}$$

2. [15 points] Consider the joint pdf $f_{X,Y}(x,y) = 6y$, for $0 < x < 1$ and $0 < y < 1-x$. [Hint: double-check your domain].

(a) Find the marginal pdf $f_X(x)$.

$$f_X(x) = \int f_{X,Y}(x,y) dy$$

$$= \int_0^{1-x} 6y dy = \left[3y^2 \right]_0^{1-x} = 3(1-x)^2$$



3 (b) Find the conditional pdf $f_{Y|X}(y)$.

$$f_{Y|X}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} \quad \begin{array}{l} \text{joint} \\ \text{marginal} \end{array} = \frac{6y}{3(1-x)^2} \quad (\text{easy})$$

$$= \frac{2y}{(1-x)^2} \quad \text{in } 0 < y < 1-x$$

$$0 < x < 1.$$

g (c) Find the covariance $\text{Cov}(X,Y)$.

$$\mu_X = E(X) = \int x f_X(x) dx$$

$$= 3 \int_0^1 x(1-x)^2 dx \quad \begin{array}{l} \text{use } \beta \text{ pdf prefactor} \\ \frac{1! 2!}{(3+1)!} = \frac{1}{12} \end{array}$$

$$= \frac{1}{4}$$

$$\mu_Y = E(Y) = \int y f_Y(y) dy \quad f_Y(y) = \int_0^{1-y} 6y dx = 6y(1-y)$$

$$= 6 \int y^2(1-y) dy = 6 \cdot \frac{2! 1!}{(3+1)!} = \frac{1}{2}$$

← Note the limits on x here defined by the domain.

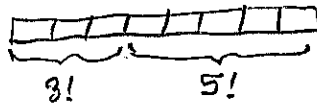
$$E(XY) = \int \int xy \cdot f_{X,Y}(x,y) dx dy = \int_0^1 \int_0^{1-y} xy \cdot 6y dx dy$$

$$= \int_0^1 6y^2 \int_0^{1-y} x dx dy = \frac{6}{2} \int_0^1 y^2(1-y)^2 dy = \frac{2! 2!}{5!} = \frac{1}{30}$$

$$\text{Cov}(X,Y) = E(XY) - \mu_X \mu_Y = \frac{1}{30} - \frac{1}{8} = \underline{\underline{-\frac{1}{40}}}$$

3. [14 points]

- 3 (a) How many ways are there of choosing a team of 3 and a team of 5 out of a class of 8?



We may permute the 8 in $8!$ ways

However we overcount by factor $3!$ for first 3 being permuted, $5!$ for the 5.

$$\Rightarrow \# \text{ ways} = \frac{8!}{3! 5!} = \binom{8}{3} = 56$$

[2-point Bonus: if the teams are both of size 4 and not distinguishable, e.g. not labelled A or B, how many ways are there now?]

If the teams are indistinguishable, i.e. team A being 1 3 5 7 and team B being 2 4 6 8 is counted as the same 'way' as A & B reversed, then the above formula overcounts by factor 2.

$$\# \text{ ways} = \frac{8!}{4! 4! 2!} = \frac{1}{2!} \binom{8}{4} = 35$$

\curvearrowright # equiv. teams

- 5 (b) A shipment of 24 eggs has 6 bad ones. You test the shipment by picking 3 at random and accepting it if at most one is bad. What is the chance that you accept the shipment?

You accept if zero or one is bad

$$\begin{aligned} \text{Chances of this are } & \frac{\binom{18}{3} \binom{6}{0}}{\binom{24}{3}} + \frac{\binom{18}{2} \binom{6}{1}}{\binom{24}{3}} \quad \} \rightarrow 2024. \\ & = \frac{1}{\binom{24}{3}} [816(1) + 53(6)] = \frac{1734}{2024} \approx 0.857 \end{aligned}$$

- 6 (c) You have 3 drawers of socks: A contains two white, B contains a white and a black, and C contains two black. However, you can't remember which drawer is which. Suppose you open a random drawer and pick out a random sock. That sock is white. What are the probabilities that the drawer you opened are A, B, or C? Please explain, using probability notation, and stating what rule of probability you used.

conditionals

$$\begin{aligned} p(W/A) &= 1 & p(K/A) &= 0 \\ p(W/B) &= \frac{1}{2} & p(K/B) &= \frac{1}{2} \\ p(W/C) &= 0 & p(K/C) &= 1 \end{aligned}$$

marginals on drawers
is uniform (random)

$$p(A) = p(B) = p(C) = \frac{1}{3}$$

Rule:

$$p(A|W) = \frac{p(W/A) p(A)}{p(W)}$$

Bayes rule

$$\left\{ \begin{aligned} &\text{with } p(W) = p(W/A)p(A) + \\ &p(W/B)p(B) + p(W/C)p(C) \\ &= \frac{1}{3}(1 + \frac{1}{2} + 0) = \frac{1}{2} \end{aligned} \right.$$

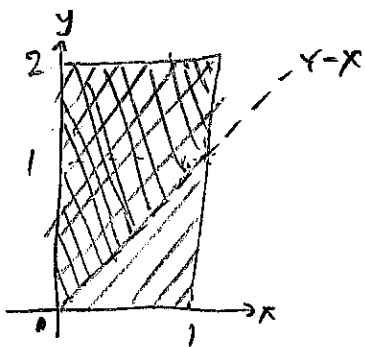
$$= \frac{1}{\frac{1}{2}} \cdot \frac{1}{3} = \frac{2}{3}$$

$$p(B|W) = \frac{p(W/B) p(B)}{p(W)} = \frac{\frac{1}{2}}{\frac{1}{2}} \cdot \frac{1}{3} = \frac{1}{3}$$

$$p(C|W) = \frac{p(W/C) p(C)}{p(W)} = 0$$

4. [18 points] Rectangular mint candies of variable size are produced by a machine. Let X and Y be random variables giving their width and length.

- 4 (a) If X has uniform pdf in the interval $[0, 1]$ and, independently, Y has uniform pdf in the interval $[0, 2]$, find the probability that $Y \geq X$.



$$\begin{aligned} f_{X,Y}(x,y) &= f_X(x) f_Y(y) = 1 \cdot \frac{1}{2} \\ &= \frac{1}{2} \text{ in } 0 \leq x \leq 1, 0 \leq y \leq 2. \end{aligned}$$

Geometric probability (area fraction) since uniform pdf.

$$p(Y \geq X) = \frac{\text{area above } y=x}{\text{total area}} = \frac{3/2}{2} = \frac{3}{4}$$

complete limits: 1 , for $0 \leq \frac{a}{x} \leq 1$, i.e. $x \geq a$.

5 (b) Given the same pdf as above, find the pdf of the candy area $A = XY$. $\frac{1}{2}$ for $0 \leq x \leq 2$.

use formula for pdf of product, since X, Y indep.

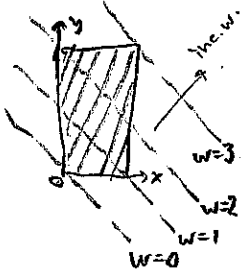
$$f_A(a) = \int \frac{1}{|x|} f_x\left(\frac{a}{x}\right) f_Y(x) dx$$

$$= \frac{1}{2} \int_a^2 \frac{1}{x} dx = \frac{1}{2} (\ln 2 - \ln a) \quad \text{for } 0 < a \leq 2$$

people found this convolution hard! $\frac{1}{2}$ for $0 \leq w-x \leq 2$ i.e. $w-2 \leq x \leq w$.

5 (c) The edge of each candy is lined by chocolate, with length given by the perimeter $P = 2(X + Y)$. Given the same pdf as above, find the pdf of the perimeter. [Hint: find pdf of $X + Y$ first, taking care to consider its various domains].

$W = X + Y$ has $f_W(w) = \int f_Y(w-x) f_X(x) dx$ 1 for $0 \leq x \leq 1$.



3 domains of w :

$0 \leq w \leq 1$: $f_W(w) = \int_0^w \frac{1}{2} dx = \frac{w}{2}$

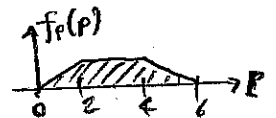
$1 \leq w \leq 2$: $f_W(w) = \int_{0,1}^1 \frac{1}{2} dx = \frac{1}{2}$

$2 \leq w \leq 3$: $f_W(w) = \int_{w-2}^1 \frac{1}{2} dx = \frac{1}{2} (1 - w + 2) = \frac{3-w}{2}$

Now $P = 2W$ so transform the pdf, $f_P(p) = \frac{1}{|2|} f_W\left(\frac{p}{2}\right)$ (linear)

so $f_P(p) = \begin{cases} \frac{p}{8} & , 0 \leq p \leq 2 \\ \frac{1}{4} & , 2 \leq p \leq 4 \\ \frac{p-6}{8} & , 4 \leq p \leq 6 \end{cases}$

0 zero otherwise.



4 (d) Finally, assume the machine is adjusted so that the candies are each square, so there is only one random variable, X , and it has uniform pdf in $[0, 1]$. Find the new pdf of the area $A = X^2$.

Now it's a nonlinear transformation of single pdf via function $A(x) = x^2$

Inverse func $A^{-1}(a) = a^{1/2}$

derivative $\frac{d}{da} A^{-1}(a) = \frac{d}{da} a^{1/2} = \frac{1}{2} a^{-1/2}$

which is monotonic in $[0, 1]$.

$f_A(a) = \left[\frac{d}{da} A^{-1}(a) \right] f_X(A^{-1}(a))$

$= \frac{1}{2} a^{-1/2} \cdot \begin{cases} 1 & \text{for } A^{-1}(a) \in [0, 1] \\ 0 & \text{otherwise} \end{cases} = \frac{1}{2} a^{-1/2}, 0 \leq a \leq 1, \text{ zero otherwise.}$

5. [25 points] Consider the model pdf $f_Y(y; \theta) = (1/\theta)e^{-y/\theta}$ for $y \geq 0$. Data $\{y_i\}$, $i = 1, \dots, n$ are collected, from which we wish to estimate θ .

5 (a) Derive the Maximum Likelihood estimator $\hat{\theta}$. Please show your working.

$$\begin{aligned} \text{Likelihood } L(\theta) &= \prod_{i=1}^n f_Y(y_i; \theta) = \frac{1}{\theta^n} \prod_{i=1}^n e^{-y_i/\theta} \\ &= \theta^{-n} e^{-\sum y_i / \theta} \end{aligned}$$

$$\ln L(\theta) = -n \ln \theta - \frac{\sum y_i}{\theta}$$

$$\frac{d}{d\theta} \ln L = -\frac{n}{\theta} + \frac{\sum y_i}{\theta^2} \quad \text{set to zero for max.}$$

$$\Rightarrow \frac{n}{\theta} = \frac{\sum y_i}{\theta^2} \quad \Rightarrow \hat{\theta} = \frac{\sum y_i}{n} = \bar{y}$$

5 (b) Compute the variance of this estimator, $\text{Var}(\hat{\theta})$, assuming the data does in fact come from the model pdf with parameter value θ .

$$\text{Var}(\hat{\theta}) = \frac{1}{n^2} \text{Var}(\sum y_i) = \frac{1}{n} \text{Var}(Y) \quad \text{since } \{y_i\} \text{ indep.}$$

$$\begin{aligned} \text{with } \text{Var}(Y) &= E(Y^2) - E(Y)^2 \\ &= \int_0^\infty y^2 \cdot \frac{1}{\theta} e^{-y/\theta} dy = \frac{\theta^3}{\theta} \int_0^\infty \left(\frac{y}{\theta}\right)^2 e^{-\frac{y}{\theta}} \frac{dy}{\theta} = \theta^2 \Gamma(3) = 2\theta^2 \end{aligned}$$

sub. $u = y/\theta$,
use gamma.

$$\text{and } E(Y) = \int_0^\infty \frac{y}{\theta} e^{-y/\theta} dy = \theta \int_0^\infty \frac{y}{\theta} e^{-y/\theta} \frac{dy}{\theta} = \theta \Gamma(2) = \theta.$$

$$\text{so } \text{Var}(\hat{\theta}) = \frac{1}{n} \cdot [2\theta^2 - (\theta)^2] = \frac{\theta^2}{n}$$

5 (c) Compare this to the Cramér-Rao lower bound. Is then $\hat{\theta}$ an efficient estimator?

CR bound : $\ln f_Y(Y; \theta) = -\ln \theta - \frac{Y}{\theta}$

expectation over rand. var Y with its pdf. $\rightarrow E \left[\frac{\partial^2}{\partial \theta^2} \ln f_Y(Y; \theta) \right] = E \left[\frac{1}{\theta^2} - \frac{2Y}{\theta^3} \right] = E \left[\frac{1}{\theta^2} \right] - \frac{2}{\theta^3} E(Y) = \frac{1}{\theta^2} - \frac{2}{\theta^2} = -\theta^{-2}$

↑ θ from above.

$$\text{CR bound: } \text{Var}(\hat{\theta}) \geq \frac{-1}{n E[-]} = \frac{\theta^2}{n} \quad \text{It is an efficient estimator.}$$

this follows from Chebyshev's Theorem.

- 3 (d) Prove that the estimator is consistent. [Hint: make sure you demonstrate everything you need to].

Asymptotically unbiased & variance vanishes as $n \rightarrow \infty$ \Rightarrow Consistent.
 \hookrightarrow true since $E(\hat{\theta}) = E(\bar{Y}) = E(Y) = \theta$, unbiased
 \hookrightarrow true since $\text{Var}(\hat{\theta}) = \frac{\text{const}}{n}$

(You don't need to integrate over the full pdf on $\hat{\theta}$)

- 5 (e) With the same assumption as before, compute the full normalized pdf of the estimator, $f_{\hat{\theta}}(u)$.
 [Hint: go back to the form of $\hat{\theta}$].

This was (intentionally) hard!

$$\hat{\theta} = \frac{1}{n} [Y_1 + Y_2 + \dots + Y_n]$$

\hookrightarrow call W , the sum of n indep. exponentially-dist. random vars.

Each $Y \sim \frac{1}{\theta} e^{-y/\theta}$ = gamma pdf with $r=1$, $\lambda = \frac{1}{\theta}$

pdf on sum of gamma variables given by summing r values (if λ 's all equal)

$\Rightarrow W \sim$ gamma with $r=n$, $\lambda = \frac{1}{\theta}$

Finally use linear transformation $\hat{\theta} = \frac{1}{n}W$ so

don't confuse u (coord) with θ (fixed val. of param).

$$f_{\hat{\theta}}(u) = n f_W(nu) = n \cdot \frac{\lambda^n}{\Gamma(n)} (nu)^{n-1} e^{-\lambda nu} = \frac{n^n}{\theta^n \Gamma(n)} \cdot u^{n-1} e^{-\frac{nu}{\theta}}$$

- 2 (f) What family is the conjugate prior, and why?

or, this is gamma with $r=n$, $\lambda = \frac{n}{\theta}$

Viewed as function of θ , $L(\theta) = c \cdot \theta^{-\text{const}} e^{-\text{const}/\theta}$

This is a new family we haven't encountered before. It is not a gamma on θ , rather the result of a gamma on $\frac{1}{\theta}$.
 Sorry for the tricky answer! - but I refuse for it to be "chug & plug".

6. [25 points] Heights of a random sample of four US males were 70, 72, 62, 68 inches.

- 5 (a) US females have an approximately normal height distribution with mean 64, variance 6. Assuming the male variance is the same as for females, use the data to test the hypothesis that μ for males is greater than that of females, at the 95% confidence level.

Sample mean $\bar{y} = \frac{70 + 72 + 62 + 68}{4} = 68$

Under $H_0: \mu = 64$,
 z statistic is $\frac{\bar{y} - \mu}{\sigma/\sqrt{n}} = \frac{68 - 64}{\sqrt{6/4}} = 3.27$

This exceeds the 1-sided critical value $z_{\alpha} = 1.64$

\Rightarrow Reject H_0 & Accept $H_1: \mu > 64$

(No continuity correction since \bar{y} is continuous variable).

- 3 (b) What is your p -value for the above test?

$p = P(\text{at least as extreme as observed})$

$= 1 - F_z(3.27) \approx 0.00048$

look up in far-right tail probability for 3.3.

- 6 (c) Instead assume the male variance σ^2 is unknown, and compute the 95% confidence interval on the mean μ for US males.

If σ^2 unknown, since $n=4$ is small, the statistic

$t = \frac{\bar{y} - \mu}{s/\sqrt{n}}$ has t -distn with $n-1 = 3$ degrees of freedom

here $s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{3} (2^2 + 4^2 + 6^2) = \frac{56}{3}$

\Rightarrow C.I. on μ is $\bar{y} \pm \frac{s}{\sqrt{n}} t_{\alpha/2, n-1}$ $\alpha = 0.05$
 $\begin{matrix} \bar{y} \\ \uparrow \\ 68 \end{matrix}$ $\underbrace{\hspace{10em}}_{2.16}$ \rightarrow look up ≈ 3.18

which gives $[61.1, 74.9]$

Note it includes now the $\mu = 64$ for females!

careful: so $\alpha = 0.2$

- 6 (d) Assuming the underlying pdf is normal, and μ is unknown, compute a 80% confidence interval on the variance σ^2 .

If μ unknown, $\frac{S^2(n-1)}{\sigma^2} \sim \chi^2$ with $n-1$ degrees of freedom.

\Rightarrow C.I. on σ^2 given by $\left[\frac{S^2(n-1)}{\chi^2_{1-\alpha/2, n-1}}, \frac{S^2(n-1)}{\chi^2_{\alpha/2, n-1}} \right]$
 $\Sigma(y_i - \bar{y})^2 = 56$
 $\left. \begin{matrix} 0.584 \\ 6.251 \end{matrix} \right\}$ using table for $df=3$, $p=0.1$ or 0.9 .

giving $[8.96, 95.9]$

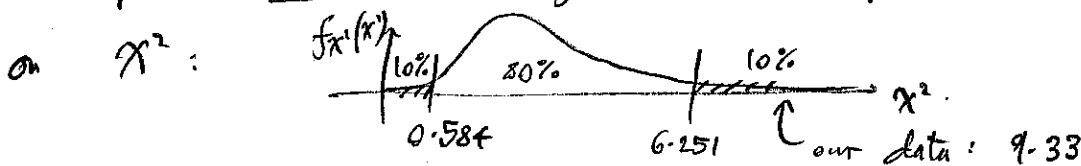
Note: on variance, not σ std. dev.

- 5 (e) Can you reject the null hypothesis that $\sigma^2 = 6$ (that of US females), at the 80% confidence level, compared against the hypothesis that $\sigma^2 \neq 6$?

Under $H_0: \sigma^2 = 6$, $\frac{S^2(n-1)}{\sigma^2} \sim \chi^2$ with $n-1$ degrees of freedom.

Our χ^2 statistic is computed to be $\frac{56}{6} = 9.33$

This falls outside the symmetric 95% probability bounds



So we reject H_0 and accept $H_1: \sigma^2 \neq 6$ at the 80% confidence level.

Note: we could have just noticed that 6 fell outside 80% C.I. from previous question.

8. [15 points] A computer LCD screen contains 1 million (that is, 10^6) pixels, each of which has an independent probability p of being 'dead' due to manufacturing defects. LCD panels are considered acceptable if they have at most 3 dead pixels.

§ (a) Say $p = 10^{-6}$. What is the average percentage of LCD panels produced that are unacceptable?

Binomial with $n = 10^6$, $p = 10^{-6}$, well approx. by Poisson
 ($n \gg 1$) ($np \sim 1$) with $\lambda = np = 1$.

$$\begin{aligned} \text{Poisson } p_x(k) &= e^{-\lambda} \frac{\lambda^k}{k!}, \quad p(\text{unacceptable}) = 1 - P(k \leq 3) = 1 - \sum_{i=0}^3 p_x(k_i) \\ &= 1 - \left(e^{-1} + e^{-1} \frac{1^1}{1!} + e^{-1} \frac{1^2}{2!} + e^{-1} \frac{1^3}{3!} \right) = 1 - \frac{8}{3} e^{-1} \approx 0.0189 \end{aligned}$$

4 (b) What is the largest p can be if the factory must produce on average at least 50% of panels which have no dead pixels? or 1.89%

$$P(\text{no dead}) = p_x(0) = e^{-\lambda} \geq 0.5$$

$$\text{take logs } \hookrightarrow -\lambda \geq \ln(0.5) = -\ln 2$$

$$\Rightarrow \lambda \leq \ln 2$$

$$\Rightarrow p \leq \frac{\ln 2}{n} = 6.93 \times 10^{-7}$$

5 (c) Quality Control examines 100 panels and counts 200 dead pixels in total. Use this data, and possibly additional assumptions, to construct a 95% confidence interval on p .

eg. if you used Bayes posterior CI, gamma.

This is a tricky question which could take you beyond the course material! (what I show you doesn't)

Let $n = 100 (10^6) = 10^8$ total pixels, $\lambda =$ Poisson parameter.

Since λ is around 200 ($\gg 1$), we may actually use normal approx to Poisson (really, same as normal approx to underlying binomial).

This takes us back to C.I. for \hat{p} from binomial data under normal approx.

$$\begin{aligned} \sigma &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \text{where } \hat{p} = \frac{k}{n} = \frac{200}{10^8} = 2 \times 10^{-6} \\ &\approx \sqrt{\frac{\hat{p}}{n}} = \sqrt{2} \cdot 10^{-7} \end{aligned}$$

$$95\% \text{ C.I. is } \hat{p} \pm 1.96\sigma = (2 \pm 0.277) \times 10^{-6} = \underline{[1.722, 2.277] \times 10^{-6}}$$

7. [15 points] Responses to the survey question, "Do you like cheese?" gave 147 out of 210 responding "Yes". (The rest responded, conveniently enough, "No").
- 6 (a) Give a 95% confidence interval on the underlying proportion p of the population that 'like' cheese.

Binomial, best estimate (MLE) is $\hat{p} = \frac{k}{n} = \frac{147}{210} = 0.7$

Since $n \gg 1$ large, may use normal approx, $\sigma = \sqrt{p(1-p)/n}$
 for pdf on \hat{p} $\approx \sqrt{\hat{p}(1-\hat{p})/n}$
 $= 0.0316$

95% C.I. is $\hat{p} \pm \overset{20.975}{1.96} \sigma = [0.638, 0.762]$

or $[63.8\%, 76.2\%]$

- 2 (b) What assumption(s) is/are needed to justify this last conclusion?

The main one is that the respondents were randomly sampled uniformly from the population.

- Also: large-sample approx. required $np \gg 1$, $n(1-p) \gg 1$ as discussed above,
 7 (c) After a brutal nationwide marketing campaign by the Cheese Board of America (no pun intended), a new survey finds 118 out of 150 responding "Yes". Test the hypothesis that the proportion has increased, against the null hypothesis that it remained the same, at the 95% confidence level. \hat{p} is close to true p .

Use normal approx. to binomial, and rules for subtracting two normal variables, $n=210$, $m=150$, $x=147$, $y=118$.

$$\begin{aligned} \text{Gives } z \text{ statistic} &= \frac{\frac{y}{m} - \frac{x}{n}}{\sqrt{\hat{p}(1-\hat{p}) \left(\frac{1}{n} + \frac{1}{m} \right)}} && \text{where under } H_0: \\ &= \frac{0.0867}{0.0471} && p_x = p_y, \\ &= 1.84 && \hat{p} = \frac{x+y}{n+m} = \frac{147+118}{210+150} \\ & && = 0.736 \end{aligned}$$

This exceeds $z_{\alpha} = 1.64$ required for 1-sided hypothesis test.

\Rightarrow Accept H_1 : p has increased from 0.7 (just).

The marketing (or something) 'worked'.