1. A study to determine the effectiveness of a drug, or serum, for the treatment of arthritis resulted in the comparison of two groups, each consisting of 400 arthritic patients. One group was inoculated with the serum, whereas the other received a placebo (an inoculation that appears to contain serum but actually is not active). After a period of time, each person in the study was asked whether their arthritic condition had improved, and the observed results are presented in the accompanying table. The question of interest is: *Do these data present evidence to indicate that the proportion of arthritic individuals who improved differs depending on whether or not they received the drug?*

Condition	Treated	Untreated
Improved	234	148
Not improved	166	252

(a) Conduct a hypothesis test using the X^2 test statistic, with $\alpha = .05$. Report (i) the null and alternative hypotheses; (ii) the expected cell counts; (iii) the test statistic; (iv) the critical value; (v) the p-value; and (vi) the conclusion.

i) $H_0: P_1 = P_2$ There is no evidence the drug is effective

$H_A: P_1 \neq P_2$ There is evidence the drug is effective.

ii) Expected Cell Counts

	Treated	Untreated	
Improved	$\frac{382 \cdot 400}{800} = 191$	191	382
x Improved	$\frac{418 \cdot 400}{800} = 209$	209	418

iii) Test Statistic

$$\frac{(234-191)^2}{191} + \frac{(148-191)^2}{191} + \frac{(166-209)}{191} + \frac{(252-209)^2}{191} = 37.05$$

IV) Critical Value

$\alpha = 0.05 \Rightarrow 3.841$

DF $(2-1)(2-1) = 1$

v) P-value : 0.000000015

vi) Since $37.05 > 3.841$ we have sufficient evidence to Reject H_0. the Drug is effective

(b) Using the Z-statistic, test the hypothesis that the proportion of treated persons who improved is equal to the proportion of untreated persons who improved, with $\alpha = .05$. *Hint:* Express each proportion as a mean. See Section 10.3 of the textbook for a refresher.

Report (i) the null and alternative hypotheses; (ii) the test statistic; (iii) the critical value; (iv) the p-value; and (v) the conclusion.

i) $H_0: P_1 = P_2$ | P_1: probability that a treated Patient improves

$H_A: P_1 \neq P_2$ | P_2: Probability that a untreated Patient improves.

ii) Test Stat

$$\frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.585 - 0.37}{\sqrt{(0.4775)-(0.5225)\left(\frac{1}{400}+\frac{1}{400}\right)}} = 6.08728$$

$$\hat{P}_1 = \frac{234}{400} \quad \hat{P}_2 = \frac{148}{400} \quad \hat{P} = \frac{234+148}{400+400} = 0.4725$$

$$\hat{q} = 1 - \hat{P} = 0.5225$$

(iii) Critical value, $\alpha = 0.05$: 1.96

IV) P-Value: 1.15×10^{-9}

Since $1.96 < 6.08728$ H_0 can be rejected

Drug is effective

the p-value, and (v) the conclusion.

(c) Prove that (assuming α is the same for both tests) the χ^2 statistic X^2 is equivalent to the square of the test statistic Z (Z^2). In other words, prove that the χ^2 test used in part (a) is equivalent to the two-tailed Z-test used in part (b).

Hint: Use the Z statistic

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}.$$

From 10.3

$$Z^2 = \frac{(\hat{P}_1 - \hat{P}_2)^2}{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \frac{n_1 n_2 (\hat{P}_1 - \hat{P}_2)^2}{(n_1 + n_2)\hat{p}\hat{q}}$$

Note that

$$\hat{p} = \frac{Y_1 + Y_2}{n_1 + n_2} = \frac{n_1\hat{P}_1 + n_2\hat{P}_2}{n_1 + n_2}$$

Given χ^2 H_0: independence of classification

H_A: dependence of classification

	Treated	Untreated	Total
improved	$C_{11} = n_1\hat{P}_1$	$C_{12} = n_2\hat{P}_2$	$C_{11} + C_{12}$
x improved	$C_{21} = n_1\hat{q}_1$	$C_{22} = n_2\hat{q}_2$	$C_{11} + C_{22}$
	$C_{11} + C_{21} = C_1$	$n_{12} + n_{22} = n_2$	$C_1 + C_2 = C$

$$\hat{E}(C_{11}) = \frac{(C_{11} + C_{12})(C_{11} + C_{21})}{C_1 + C_2} = \frac{(Y_1 + Y_2)(C_{11} + C_{21})}{C_1 + C_2}$$

$$= C_1\hat{P}$$

So $\hat{E}(C_{21}) = C_1\hat{q}$ $\hat{E}(C_{12}) = C_2\hat{P}$

$$X^2 = \frac{c_1^2(\hat{p}_1 - \hat{p})^2}{c_1\hat{p}} + \frac{c_1^2(\hat{q}_1 - \hat{q})^2}{c_1\hat{q}} + \frac{c_2^2(\hat{p}_2 - \hat{p})^2}{c_2\hat{p}} + \frac{c_2^2(\hat{q}_2 - \hat{q})^2}{c_2\hat{q}}$$

$$X^2 = \frac{c_1(\hat{p}_1 - \hat{p})^2}{\hat{p}} + \frac{c_1[(1-\hat{p}_1) - (1-\hat{p})]^2}{\hat{q}} + \frac{c_1(\hat{p}_2 - \hat{p})^2}{\hat{p}} + \frac{c_2[(1-\hat{p}_2) - (1-\hat{p})]^2}{\hat{q}}$$

$$X^2 = \frac{c_1(\hat{p}_1 - \hat{p})^2}{\hat{p}\hat{q}} + \frac{c_2(\hat{p}_2 - \hat{p})^2}{\hat{p}\hat{q}}$$

simplifies to

$$X^2 = \frac{c_1}{\hat{p}\hat{q}}\left(\frac{c_1\hat{p}_1 + c_2\hat{p}_1 - c_1\hat{p}_1 - c_2\hat{p}_2}{c_1 + c_2}\right) + \frac{c_1}{\hat{p}\hat{q}}\left(\frac{c_1\hat{p}_2 + c_1\hat{p}_2 - c_1\hat{p}_1 - c_2\hat{p}_2}{c_1 + c_2}\right)^2$$

$$= \frac{c_1c_2(\hat{p}_1 - \hat{p}_2)^2}{\hat{p}\hat{q}(c_1 + c_2)} \Rightarrow$$ which shows the test are equivalent.

2. Consider the following model for the responses measured in a randomized block design containing b blocks and k treatments:

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij},$$

where:

Y_{ij} = response to treatment i in block j;

μ = overall mean;

τ_i = nonrandom effect of treatment i, where $\sum_{i=1}^{k} \tau_i = 0$;

β_j = random effect of block j, where β_js are independent, normally distributed random variables with $E[\beta_j] = 0$ and $V(\beta_j) = \sigma_B^2$ for $j = 1, 2, ..., b$;

ϵ_{ij} = random error terms where ϵ_{ij}s are independent, normally distributed random variables with $E[\epsilon_{ij}] = 0$ and $V(\epsilon_{ij}) = \sigma_\epsilon^2$ for $i = 1, 2, ..., k$ and $j = 1, 2, ..., b$.

$Cov(X, Y) = 0$ iff $X \perp\!\!\!\perp Y$, Same w/ Y_{ij} and $Y_{ij'}$

$Cov(\mu + \tau_i + \beta_{j'} + \varepsilon_{ij'}, \mu + \tau_i + \beta_j + \varepsilon_{ij})$, μ and τ_i are constant

$= Cov(\beta_{j'} + \varepsilon_{ij'}, \beta_j + \varepsilon_j)$

$= Cov(\beta_j, \beta_{j'}) + Cov(\varepsilon_{ij'}, \beta_j) + Cov(\beta_{j'}, \varepsilon_{ij}) + Cov(\varepsilon_{ij'}, \varepsilon_{ij})$

We know $\beta_j \perp\!\!\!\perp \varepsilon_{ij}$, and $\beta_j \cdot Y \overset{iid}{\sim} \varepsilon_{ij} \cdot Y$

$= 0 + 0 + 0 + 0 = 0$

Same w/ $Y_{i'j'}$ and Y_{ij}

$Cov(\mu + \tau_{i'} + \beta_{j'} + \varepsilon_{i'j'}, \mu + \tau_i + \beta_j + \varepsilon_{ij})$

$= Cov(\beta_{j'} + \varepsilon_{i'j'}, \beta_j + \varepsilon_{ij})$

$= Cov(\beta_j, \beta_{j'}) + Cov(\varepsilon_{i'j'}, \beta_j) + Cov(\beta_{j'}, \varepsilon_{ij}) + Cov(\varepsilon_{i'j'}, \varepsilon_{ij})$

B/C $\beta_j \perp\!\!\!\perp \varepsilon_{ij}$ $\beta_j Y \overset{iid}{\sim} \varepsilon_{ij} Y$

$= 0 + 0 + 0 + 0 = 0$

b). $Cov(Y_{i'j}, Y_{ij})$

$= Cov(\mu + \tau_{i'} + \beta_j + \varepsilon_{i'j},\ \mu + \tau_i + \beta_i + \varepsilon_{ij})$

$= Cov(\beta_j + \varepsilon_{ij},\ \beta_j + \varepsilon_{i'j})$

$= Cov(\beta_j, \beta_j) + Cov(\varepsilon_{i'j}, \beta_j) + Cov(\beta_j, \varepsilon_{ij}) + Cov(\varepsilon_{i'j}, \varepsilon_{ij})$

$= V(\beta_j) + 0 + 0 + 0 = \sigma_\beta^2$

c). if $\sigma_\beta^2 = Var(\beta_j) = 0$, then $Y_{i'j} \perp\!\!\!\perp Y_{ij}$, $Cov(Y_{i'j}, Y_{ij}) = 0$

d) $E(Y_{ij}) = E(\mu + \tau_i + \beta_j + \varepsilon_{ij})$

$= E(\mu) + E(\tau_i) + E(\beta_j) + E(\varepsilon_{ij})$

$= \mu + \tau_i + 0 + 0$

$= \mu + \tau_i$

$Var(Y_{ij}) = Var(\mu + \tau_i + \beta_j + \varepsilon_{ij})$

$= var(\beta_j + \varepsilon_{ij})$

$= Var(\beta_j) + Var(\varepsilon_{ij}) + 2\,Cov(\beta_j, \varepsilon_{ij})$

$= \sigma_\beta^2 + \sigma_\varepsilon^2 + 0 = \sigma_\beta^2 + \sigma_\varepsilon^2$

e) $E(\bar{Y}_{i*}) = E\left(\frac{1}{b}\sum_{j=1}^{b} Y_{ij}\right) = \frac{1}{b} E\left(\sum_{j=1} b Y_{ij}\right)$

$= \frac{1}{b} \cdot b\, E(Y_{ij})$

$= E(Y_{ij})$

$= \mu + \tau_i \Rightarrow \bar{Y}_{i*}$ is unbiased Estimator

$$
\begin{aligned}
V(\bar{Y}_{i*}) &= V\left(\frac{1}{b}\sum_{j=1}^{b} Y_{i,j}\right) \\
&= \frac{1}{b^2} V\left(\sum_{j=1} b\, Y_{i,j}\right) \\
&= \frac{1}{b^2} \cdot b\, V(Y_{i,j}) \\
&= \frac{1}{b}\left(\sigma_\beta^2 + \sigma_\varepsilon^2\right)
\end{aligned}
$$

f) Since we have $E(\bar{Y}_{i*}) = \boxed{\mu + \tau_i}$ from part e).

and $E(Y_{i,j}) = \boxed{\mu + \tau_i}$ from part d).

It's an unbiased estimator of the mean response to treatment i.

3. For a comparison of the academic effectiveness of two junior high schools A and B, an experiment was designed using ten sets of identical twins, where each twin had just completed the sixth grade. In each case, the twins in the same set had obtained their previous schooling in the same classrooms at each grade level. One child was selected at random from each set and assigned to school A. The other was sent to school B. Near the end of the ninth grade, an achievement test was given to each child in the experiment. The results are shown in the accompanying table.

Twin Pair	A	B
1	67	39
2	80	75
3	65	69
4	70	55
5	86	74
6	50	52
7	63	56
8	81	72
9	86	89
10	60	47

(a) Using the sign test, test the hypothesis that the two schools are the same in academic effectiveness, as measured by scores on the achievement test, against the alternative that the schools are not equally effective. What would you conclude with $\alpha = .05$?

A_i	B_i	$D_i = A_i - B_i$
67	39	28
80	75	5
65	69	-4
70	55	15
86	74	8
50	52	-2
63	56	7
81	72	9
86	89	-3
60	47	13

$H_0: F(x) = G(y) \iff p = \frac{1}{2}$

$H_A: F(x) \neq G(y) \iff p \neq \frac{1}{2}$

$M = 7 \qquad n - M = 3$

$$P(M \geq 7 \mid p = 0.5) = \binom{10}{7}(0.5)^{10} + \binom{10}{8}(0.5)^{10} + \binom{10}{9}(0.5)^{10} + \binom{10}{10}(0.5)^{10}$$

$= 0.171875$

$P(M \leq 3 \mid p = 0.5) \cdot 2 = 0.3438$

$\alpha = 0.05 <$ P-value $= 0.3438$

Fail to Reject H_0

(b) Suppose it is suspected that junior high school A has a superior faculty and better learning facilities. Test the hypothesis of equal academic effectiveness against the alternative that school A is superior. What is the p-value associated with this test?

$H_0: P = \frac{1}{2}$ $\quad H_A: P > \frac{1}{2}$

P-Value $= P(M \geq 7 \mid P = 0.05) = 0.171875$ one tailed.

(c) Repeat the test in (a), using the Wilcoxon signed-rank test. Compare your answers.

A_i	B_i	$D_i = A_i - B_i$	Rank
67	39	28	10
80	75	5	4
65	69	-4	3
70	55	15	9
86	74	12	6
50	52	-2	1
63	56	7	5
81	72	9	7
86	89	-3	2
60	47	13	8

$H_0: f(x) = g(x)$

$H_a: f(x) \neq g(x)$

$T^+ = |D_i^+| = 10+4+9+6+5+7+8 = 49$

$T^- = |D_i^-| = 3+1+2 = 6$

$\min\{T^+, T^-\} = 6$

With $\alpha = 0.05$, Critical Value is 8.

Since $6 < 8$, Reject H_0

Different conclusion w/ (a).

4. Let $Y_1, Y_2, ..., Y_n$ denote a random sample from an exponentially distributed population with density $f(y|\theta) = \theta e^{-\theta y}$, $0 < y$. (Note that the mean of this population is $\mu = \frac{1}{\theta}$.) Use the conjugate gamma (α, β) prior for θ to find the following:

   (a) The joint density, or $f(y_1, y_2, ..., y_n, \theta)$;

   (b) The marginal density, or $m(y_1, y_2, ..., y_n)$;

   (c) The posterior density for $\theta | (y_1, y_2, ..., y_n)$.

$\frac{\text{Exp is special case of Gamma}}{\text{which } \alpha = 1}$

a) By Def in ch 16.2

$$f(y_1, y_2, y_3 \cdots y_n, \theta) = L(y_1 \cdots y_n | \theta) \times g(\theta)$$

Joint Density is given by

$$f(y_1, y_2, \cdots y_n, \theta) = \prod_{i=1}^{n} (\theta e^{-\theta y_i}) \times \frac{1}{\Gamma(\alpha)\beta^\alpha} \theta^{\alpha-1} e^{-\frac{\theta}{\beta}}$$

$$= \frac{\theta^{n+\alpha-1}}{\Gamma(\alpha)\beta^\alpha} \cdot e^{-\theta \sum_{i=1}^{n} y_i - \frac{\theta}{\beta}}$$

$$\boxed{= \frac{\theta^{n+\alpha-1}}{\Gamma(\alpha)\beta^\alpha} \exp\left(\frac{-\theta}{\frac{\beta}{\beta \sum_{i=1}^{n} y_i + 1}}\right)} \quad \square$$

b) Given the Answer obtained in a)
And by definition of Marginal Density

$$m(y_1 \ldots, y_n) = \int_0^\infty f(y_1, \cdots y_n, \theta) \, d\theta$$

$$= \int_0^\infty \frac{\theta^{n+\alpha-1}}{\Gamma(\alpha)\beta^\alpha} \exp\left(-\theta / \frac{\beta}{\beta \sum_{i=1}^{n} y_i + 1}\right)$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty \theta^{n+\alpha-1} \exp\left(-\theta / \frac{\beta}{\beta \sum_{i=1}^{n} y_i + 1}\right)$$

$$\boxed{= \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \Gamma(n+\alpha) \left(\frac{\beta}{\beta \sum_{i=1}^{n} y_i + 1} \right)^{n+\alpha}}$$

Note:

B/C Density of any dist $=1$, So is Gamma.

$$\int_0^{\infty} \frac{1}{\Gamma(n+\alpha)\left(\frac{\beta}{\beta \sum_{i=1}^{n} y_i + 1}\right)^{n+\alpha}} \cdot x^{n+\alpha} \cdot \exp\left(-x \Big/ \frac{\beta}{\beta \sum_{i=1}^{n} y_i + 1}\right) dx = 1$$

c) $$g^{*}(\theta \mid y_1 \ldots y_n) = \frac{f(y_1, \ldots, y_n, \theta)}{m(y_1 \ldots y_n)}$$

$$\boxed{= \frac{1}{\Gamma(n+\alpha)\left(\frac{\beta}{\beta \sum_{i=1}^{n} y_i + 1}\right)^{n+\alpha}} \theta^{n+\alpha-1} \exp\left(-\theta \Big/ \frac{\beta}{\beta \sum_{i=1}^{n} y_i + 1}\right)}$$