Question 1: Chapter 15 #1

Prove:

1. $Y \perp Z$.
2. $\psi = 1$.
3. $\gamma = 0$.
4. For $i, j \in \{0, 1\}$, $p_{ij} = p_{i\cdot} p_{\cdot j}$.

Y and Z are independent if and only if $P(Y = j, Z = i) = P(Y = j)P(Z = i)$ for all i, j. Therefore, $p_{ij} = p_{i\cdot} p_{\cdot j}$ $\forall i, j \in 0, 1$. Thus, 1 and 4 are equivelant.

For 2, the odds ratio ψ can be defined as, for two events E and D, $\psi = \frac{odds(D|E)}{odds(D|E^C)}$, where $odds(D|E) = \frac{P(D|E}{1-P(D|E)}$

$$\psi = 1 \implies \frac{odds(Y|Z=1)}{odds(Y|Z=0)} = 1 \tag{1}$$

$$\implies odds(Y|Z=1) = odds(Y|Z=0) \tag{2}$$

$$\frac{P(Y|Z=1)}{1 - P(Y|Z=1)} = \frac{P(Y|Z=0)}{1 - P(Y|Z=0)} \tag{3}$$

$$P(Y|Z=0)(1 - P(Y|Z=1)) = P(Y|Z=1)(1 - P(Y|Z=0)) \tag{4}$$

$$P(Y|Z=0) - P(Y|Z=0)P(Y|Z=1) = P(Y|Z=1) - P(Y|Z=1)P(Y|Z=0) \tag{5}$$

$$P(Y|Z=0) = P(Y|Z=1) \tag{6}$$

Therefore, Y is independent of Z as its probability is the same for all values that Z can take. Thus, (1) and (2) are equivelant.

(3) follows exactly as $\gamma = \log(\psi)$ and $\psi = 1$ so $\gamma = \log(1) = 0$.

Thus, (1) (2) (3) and (4) are equivelant

Question 2: Chapter 15 #4

The *New York Times* (January 8, 2003, page A12) reported the following data on death sentencing and race, from a study in Maryland:

	Death Sentence	**No Death Sentence**
Black Victim	14	641
White Victim	62	594

Analyze the data using the tools from this chapter. Interpret the results. Explain why, based only on this information, you can't make causal conclusions. (The authors of the study did use much more information in their full report.)

Answer:

This problem can be represented as two binary random variables. We can say that the random variable $Y = 0$ if a death sentence is issued and $Y = 1$ if a death sentence is not issued and the random variable $Z = 0$ if the victim is black and $Z = 1$ if the victim is white.

Let $X = (X_{00}, X_{01}, X_{10}, X_{11})$ be the vector of counts as in the table below (where Y is the random variable for death sentence and Z for race)

	$Y = 0$	$Y = 1$	
$Z = 0$	X_{00}	X_{01}	$X_{0\cdot}$
$Z = 1$	X_{10}	X_{11}	$X_{1\cdot}$
	$X_{\cdot 0}$	$X_{\cdot 1}$	$n = X_{\cdot\cdot}$

We are interested in measuring the vector of counts X. We know that each entry in X has a certain probability of occuring, which we can denote $p_{ij} = \mathbb{P}(Y = i, Z = j)$ and the vector $p = (p_{00}, p_{01}, p_{10}, p_{11})$

Thus, as X is a vector of counts, distrubted accordingly by the vector p, we can say $X \sim Multinomial(n, p)$. We can now use Pearson's χ^2 test statistic for independence:

$$U = \sum_{i=0}^{1} \sum_{j=0}^{1} \frac{(X_{ij} - E_{ij})^2}{E_{ij}}$$

Where, $E_{ij} = \frac{X_{\cdot j} X_{i \cdot}}{n}$

Filling in the data for the count table, we get:

	$Y = 0$	$Y = 1$	
$Z = 0$	14	641	655
$Z = 1$	62	594	656
	76	1235	1311

Computing U we get:

$$U = \frac{(14 - \frac{655*76}{1311})^2}{\frac{655*76}{1311}} + \frac{(641 - \frac{655*1235}{1311})^2}{\frac{655*1235}{1311}} + \frac{(62 - \frac{76*656}{1311})^2}{\frac{76*656}{1311}} + \frac{(594 - \frac{1235*656}{1311})^2}{\frac{1235*656}{1311}}$$
$$U = 32.104$$

If we use this statistic in Pearson χ^2 test with confidence $\alpha = 0.05$, as we know that this statistic is distributed as a χ_1^2. We can use the value for χ_1^2 at $\alpha = 0.05$. This is 3.841. $|U| > 3.841$ as $32.104 > 3.841$, so we reject H_0 which is that there is no correlation between race and death sentencing.

The p-value for this is $2(1 - \Phi(|U|)$ where Φ is the CDF for a χ_1^2. When we compute this we get that the p-value is $2.922 \cdot 10^{-8}$. A p-value so small is quite substantial evidence to reject H_0.

Therefore, we know that there is a correlation between the data for the death sentencing rate and the race of the prisoner. This is not enough to imply a causal relationship between the two, as there could be many other causes. $\square$

Question 3: Supplemental

3.1 Independent implies uncorrelated

Show that if (X, Y) are independent, then they are uncorrelated.
Proof:

If (X, Y) are independent, then $\mathbb{P}(X = x|Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$

$$\mathbb{E}(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \cdot \mathbb{P}(X = x|Y = y)dydx \tag{7}$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \cdot \mathbb{P}(X = x)\mathbb{P}(Y = y)dydx \tag{8}$$
$$= \int_{-\infty}^{\infty} x \cdot \mathbb{P}(X = x)dx \int_{-\infty}^{\infty} y \cdot \mathbb{P}(Y = y)dy \tag{9}$$
$$= \mathbb{E}(X)\mathbb{E}(Y) \tag{10}$$

Therefore if X and Y are independent, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$; they are uncorrelated.

3.2 Uncorrelated yet dependent

Let X random variable be such that $P(X = 1) = 1/2$ and $P(X = -1) = 1/2$. Let $Y = X^2$.

$$\mathbb{E}(XY) = \mathbb{E}(X \cdot X^2) \tag{11}$$
$$= \mathbb{E}(X^3) = \mathbb{E}(X) \qquad \text{as } X = X^3 \tag{12}$$
$$= 0 \qquad \because P(X = 1) = 1/2 \; and \; P(X = -1) = 1/2 \tag{13}$$

As $\mathbb{E}(XY) = 0$, X and Y are uncorrelated by definition, but Y is entirely dependent on X

3.3 Uncorrelated Joint Gaussians

Let X and Y be jointly Gaussian random variables, which means the joint distribution of (X, Y) is a bivariate normal distribution, fully described by:

- Their means $\mu_X = \mathbb{E}[X]$ and $\mu_Y = \mathbb{E}[Y]$,
- Their variances $\sigma_X^2 = \mathrm{Var}(X)$ and $\sigma_Y^2 = \mathrm{Var}(Y)$,
- Their covariance $\mathrm{Cov}(X, Y) = \rho\sigma_X\sigma_Y$, where ρ is the correlation coefficient.

X and Y are **uncorrelated**, which means:

$$\mathrm{Cov}(X, Y) = 0.$$

which implies that $\rho = 0$, as one of ρ, σ_X, σ_Y has to be 0 (and we know that the variance of either random variable is trivially not 0).

Proof

Since (X, Y) is jointly Gaussian, the joint distribution of (X, Y) is characterized by the following bivariate normal density:

$$f_{X,Y}(x, y) =$$
$$\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right)\right)$$

Substituting $\rho = 0$ into the joint density function, we get:

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y} \exp\left(-\frac{1}{2}\left(\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right)\right).$$

This expression factors as:

$$f_{X,Y}(x, y) = \left(\frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right)\right) \times \left(\frac{1}{\sqrt{2\pi\sigma_Y^2}} \exp\left(-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right)\right).$$

This can be rewritten as:

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y),$$

where $f_X(x)$ and $f_Y(y)$ are the marginal probability density functions of X and Y, respectively.

Conclusion:

The factorization $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$ shows that the joint distribution of (X,Y) is the product of the marginals of X and Y. This factorization implies that X and Y are independent, since the joint density is exactly the product of the marginal densities. We know that for all random variables that independence implies uncorrelated, but we have just proven that for joint gaussians, independence of X and Y implies uncorrelated and uncorrelated X and Y implies independence of X and Y. This is a special case of (b) above as it is the if and only if condition that is true in this special case, but is not true in general as we showed in (b).