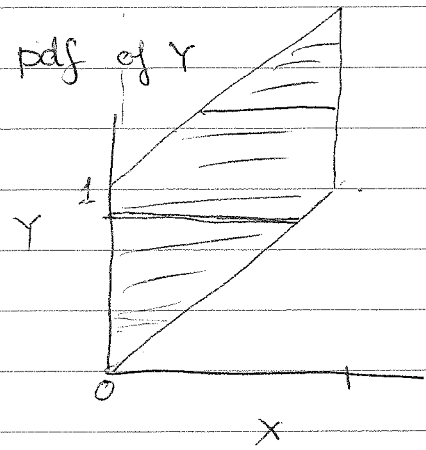
STAT 7203 - Assignment 3 solutions

Q1 (a) The joint pdf of (X,Y) is

$$f_{X,Y}(x,y) = f_{Y|X}(y|x) * f_X(x)$$

$$= \begin{cases} 1, & x \in [0,1], \ y \in [x, 1+x] \\ 0, & \text{else} \end{cases}$$

We integrate out X to get the marginal pdf of Y



$$f_Y(y) = \int f_{X,Y}(x,y)\, dx$$

$$= \int_{(0 \vee y-1)}^{(y \wedge 1)} 1 \, dx$$

$$= \begin{cases} \int_0^y 1\, dx = y & y \in [0,1] \\ \int_{y-1}^1 1\, dx = 2-y & y \in [1,2] \end{cases}$$

(b) $\mathrm{cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

$$\mathbb{E}[XY] = \mathbb{E}[X\, \mathbb{E}[Y|X]]$$

$$\mathbb{E}[Y|X=x] = \int_x^{1+x} y \cdot 1 \, dy = \frac{1}{2} y^2 \Big|_x^{1+x} = \frac{1}{2}\left((1+x)^2 - x^2\right)$$

$$= \frac{1}{2}(1+2x)$$

$$\mathbb{E}[XY] = \mathbb{E}\left[\frac{1}{2}X(1+2X)\right] = \frac{1}{2}\mathbb{E}[X] + \mathbb{E}[X^2]$$

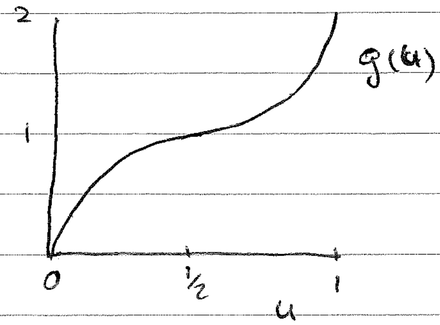
As $X \sim U[0,1]$ $\quad \mathbb{E}[X] = \int_0^1 x\, dx = \frac{1}{2} \quad \mathbb{E}[X^2] = \int_0^1 x^2\, dx = \frac{1}{3}$

So $\mathbb{E}[XY] = \frac{1}{4} + \frac{1}{3} = 7/12$

Finally, $\mathbb{E}[Y] = \mathbb{E}[\,\mathbb{E}[Y|X]\,] = \mathbb{E}[\tfrac{1}{2}(1+2X)] = \tfrac{1}{2} + \mathbb{E}[X] = 1$

so: $\operatorname{cov}(X,Y) = 7/12 - \tfrac{1}{2}\times 1 = 1/12$

(c) Inverse of g:



For $z \in [0,1]$

$$g(u) = z$$
$$\sqrt{2u} = z$$
$$\Rightarrow \quad u = z^2/2$$

for $z \in [1,2]$

$$g(u) = z$$
$$2 - \sqrt{2(1-u)} = z$$
$$2 - z = \sqrt{2(1-u)}$$
$$u = 1 - \tfrac{1}{2}(2-z)^2$$

$$F_Z(z) = \mathbb{P}(g(U) \leq z)$$
$$= \mathbb{P}(U \leq g^{-1}(z))$$
$$= \begin{cases} \tfrac{1}{2}z^2, & z \in (0,1) \\ 1 - \tfrac{1}{2}(2-z)^2, & z \in [1,2) \end{cases}$$

The pdf of Z is

$$f_Z(z) = \frac{d}{dz} F_Z(z)$$

$$= \begin{cases} z, & z \in (0,1) \\ 2 - z, & z \in (1,2) \end{cases}$$

This is the same as the pdf of Y in part (a).

Q2 (a) Test H_0: 'Training method' and 'accuracy' are independent.
against H_1: some association between 'training method' and 'accuracy'.

Expected counts

card-sort	$\frac{58\times44}{91} = 28.0$	$\frac{20\times44}{91} = 9.7$	$\frac{13\times44}{91} = 6.3$	44
triage-trainer	$\frac{58\times47}{91} = 30.0$	$\frac{20\times47}{91} = 10.3$	$\frac{13\times47}{91} = 6.7$	47
	58	20	13	91

$$X^2 = \frac{(28.0-24)^2}{28.0} + \frac{(30.0-34)^2}{30.0} + \frac{(9.7-11)^2}{9.7} + \frac{(10.3-9)^2}{10.3} + \frac{(6.3-9)^2}{6.3}$$

$$+ \frac{(4-6.7)^2}{2} = 3.75$$

$df = (\text{rows}-1)\times(\text{colums}-1)$
$= (2-1)\times(3-1) = 2.$

Compare X^2 to a χ^2_2 distribution to get p-value

$\mathbb{P}(\chi^2_2 \geq 2.773) = 0.25$ and $\mathbb{P}(\chi^2_2 \geq 4.605) = 0.1$

The p-value is between 0.1 and 0.25. This is no evidence/ inconclusive evidence against the null hypothesis, suggesting accuracy and training method are independent.

(b) let p_c be the proportion who get 8/8 from card-sort group
" p_T " " " " " " " triage-trainer " .

A 95% CI for $p_c - p_T$ is:

$\hat{p}_c = \frac{24}{44} \doteq 0.54$, $n_c = 44$ $\hat{p}_T = \frac{34}{47} \simeq 0.72$ $n_T = 47$

$$(\hat{p}_C - \hat{p}_T) \pm 1.96 \sqrt{\frac{\hat{p}_C(1-\hat{p}_C)}{n_C} + \frac{\hat{p}_T(1-\hat{p}_T)}{n_T}}$$

$$-0.178 \pm 1.96 \sqrt{\frac{0.54 \times 0.45}{44} + \frac{0.72 \times 0.28}{47}}$$

$$-0.178 \pm 0.195$$

(c) let μ_C be the mean time taken to triage casualties by person trained using card-sort method and
let μ_T be ... trained using triage-trainer

Test $H_0 : \mu_C = \mu_T$ against $\mu_C \neq \mu_T$

test statistic $t = \dfrac{\text{estimate} - \text{hypothesised}}{\text{s.e.(estimate)}}$

pooled sample variance $s_p^2 = \dfrac{(44-1) \times 74^2 + (47-1) \times 62^2}{44+47-2}$

$= 4632.5$

test statistic $= \dfrac{(435-456) - 0}{\sqrt{4632.5}\sqrt{1/44 + 1/47}}$

$= -1.471$

p-value $= 2 \times \min\{\mathbb{P}(T_{89} \geq -1.471), \mathbb{P}(T_{89} \leq -1.471)\}$

$= 2 \times \mathbb{P}(T_{89} \leq -1.471)$

$\mathbb{P}(T_{80} \leq -1.292) = 0.1$ and $\mathbb{P}(T_{80} \leq -1.664) = 0.05$

so the p-value is between 0.1 and 0.2, this is inconclusive evidence / no evidence against the null hypothesis, suggesting the mean time taken to triage the eight casualties is the same for the two training methods.

(d) We want a 95% CI for μ_T

$$\text{estimate} \pm t_{46;0.975} \times \text{s.e.(estimate)}$$

$$456 \pm 2.021 \times \frac{62}{\sqrt{47}}$$

(using 40 degrees of freedom from tables, Actual value using MATLAB is 2.013)

$$456 \pm 18.3 \text{ (s)}.$$