جامعة بنها – كلية التربية　　　　بسم الله الرحمن الرحيم　　　　الفرقة الثالثة عام (لائحة جديدة) : شعبة رياضيات

قسم الرياضيات　　　　امتحان نهاية الفصل الدراسي الثاني لعام2012 /2013 م　مادة (إحصاء)　　　　الزمن : ساعاتان

إجابة أمتحان

الفرقة : ثالثة رياضيات شعبة الرياضيات　　　　المادة : إحصاء

يوم الأمتحان : الأربعاء 29/ 5 / 2013 م ورقة كاملة

أستاذ المادة : أ . د . / حسني كامل عبد المقصود أستاذ متفرغ بكلية العلوم جامعة بنها

إجابة السؤال الأول (أ):

من الجدول واضح ان احتمال ان يكون عدد الحشرات من النوع B أكبر من عدد الحشرات من النوع A هو :

$$P(Y > X) = P(Y = 2, X = 1) = 0.20$$

وكذلك يمكن الحصول علي التوزيعات الهامشية لكلا المتغيرين وهما

X	1	2	3	4
$p(x_i)$	0.28	0.32	0.20	0.20

Y	0	1	2
$q(y_j)$	0.20	0.43	0.37

$$\mu_X = (1)(.28) + (2)(.32) + (3)(.20) + (4)(.20) = 2.32$$

$$E[X^2] = (1)^2(.28) + (2)^2(.32) + (3)^2(.20) + (4)^2(.20) = 6.56$$

$$\sigma^2{}_X = E[X^2] - [\mu_X]^2 = 1.18 \rightarrow \sigma_X = 1.086$$

$$\mu_Y = 1.17$$

$$\sigma^2{}_Y = 0.54 \rightarrow \sigma_Y = 0.735$$

ومن جدول التوزيع المشترك يمكن حساب $E[XY]$ كالتالي :

$$E[XY] = (1*1)(0.8) + (1*2)(0.15) + (1*3)(0.10) + (1*4)(0.10) + (2*1)(0.20)$$
$$+ (2*2)(0.12) + (2*3)(0.05) + (2*4)(0) = 2.26$$
$$\therefore Cov[X, Y] = E[XY] - \mu_X\mu_Y = 2.26 - 2.714 = -0.454$$

وبالتالي فإن :

$$\rho(X, Y) = \frac{Cov[X, Y]}{\sigma_X\sigma_Y} = \frac{-0.454}{(1.09)(0.735)} = -0.561$$

من جدول التوزيع المشترك يمكن الحصول علي توزيع الاحتمالي للمتغير T كالتالي :

T	2	3	4	5
$p(t_i)$	0.13	0.40	0.32	0.15

و منة يمكن حساب

$$E[T] = 3.49 \quad , \quad E[T^2] = 12.99 \quad , \quad \sigma^2{}_T = 0.81$$

وبسهولة يمكن التحقق من النتائج المطلوبة .

إجابة السؤال الأول (ب):

إذا كانت $G(y)$ ترمز إلي دالة التوزيع التراكمية للمتغير Y عند النقطة y . فإن

$$
\begin{aligned}
G(y) &= P(Y \le y) \\
&= P(X^3 \le y) \\
&= P\left(X \le y^{1/3} \right) \\
&= \int_0^{y^{\frac{1}{3}}} \frac{1}{2} x dx \\
&= \frac{1}{2}\left[\frac{x^2}{2}\right]_0^{y^{\frac{1}{3}}} = \frac{1}{4} y^{\frac{2}{3}}
\end{aligned}
$$

وبالتالي فإن $g(y) = \frac{1}{6} y^{-\frac{1}{3}}$ لكل $0 < y < 8$ ، $g(y) = 0$ بخلاف ذلك .

إجابة السؤال الثاني (أ):

يمكن إيجاد دالة الكثافة الهامشية لكل منهما وهي

$$f(x) = \int_0^1 4xydy = 2x \qquad ; \qquad 0 \le x \le 1$$

$$g(y) = \int_0^2 4xydx = 2y \qquad ; \qquad 0 \le y \le 1$$

$$E[X] = \mu_X = \int xf(x)dx = \int_0^1 x \times 2xdx = \frac{2}{3}$$

$$E[Y] = \mu_Y = \int yg(y)dy = \int_0^1 2y \times ydy = \frac{2}{3}$$

$$E[X^2] = \int x^2 f(x)dx = \int_0^1 x^2 \times 2xdx = \frac{1}{2}$$

$$E[Y^2] = \int y^2 g(y)dy = \int_0^1 y^2 \times 2ydy = \frac{1}{2}$$

$$Var[X] = \sigma_X^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

$$Var[Y] = \sigma_Y^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

$$E[XY] = \int\int xyf(x,y)dxdy = \int_0^1\int_0^1 xy \times 4xydxdy = \frac{4}{9}$$

$$Cov[X,Y] = E[XY] - E[X]E[Y] = \frac{4}{9} - \frac{2}{3} \times \frac{2}{3} = 0$$

$$\rho[X,Y] = \frac{Cov[X,Y]}{\sigma_X \sigma_Y} = 0 \Big/ \sqrt{\frac{1}{18} \times \frac{1}{18}} = 0$$

$$P(0 < X < 0.4; 0.5 < Y < 0.8) = \int_0^{0.4} dx \int_{0.5}^{0.8} 4xydy = \left[x^2\right]_0^{0.4} \left[y^2\right]_{0.5}^{0.8} = 0.0624$$

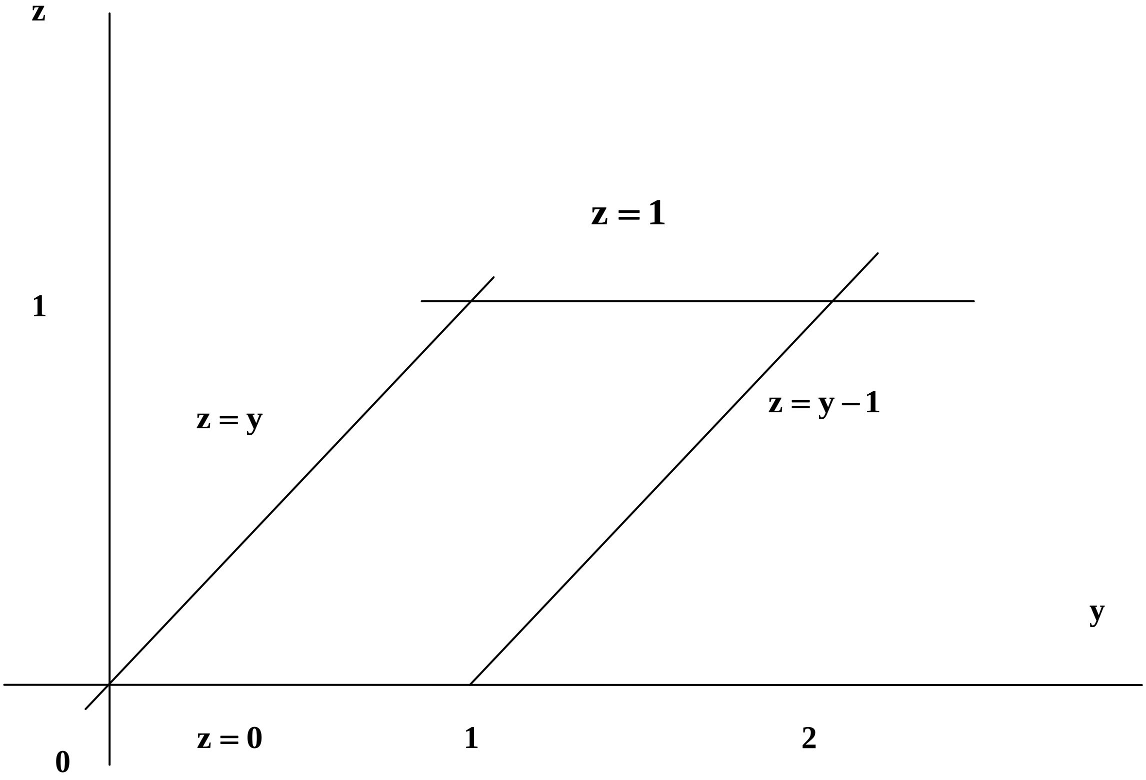
$$P(0.2 < X < 0.6) = \int_{0.2}^{0.6} f(x)dx = \int_{0.2}^{0.6} 2xdx = 0.32$$

إجابة السؤال الثاني (ب):

(I) . بحل كل من المعادلتين $y = x_1 + x_2$ ، $z = x_2$ للقيم x_1 ، x_2 فإننا نجد أن $x_1 = y - z$ ، $x_2 = z$ ومن هذه العلاقات نحصل علي

$$J = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1$$

وحيث أن التحويل تناظر واحد إلي واحد فأن الدوال تنقل المنطقة $1 > x_1 > 0$ ، $1 > x_2 > 0$ في المستوي $x_1 x_2$ إلي المنطقة $z < y < z + 1$ في المستوي yz



وباستخدام $g(y_1, y_2) = f[W_1(y_1, y_2), W_2(y_1, y_2)].|J|$ نجد أن

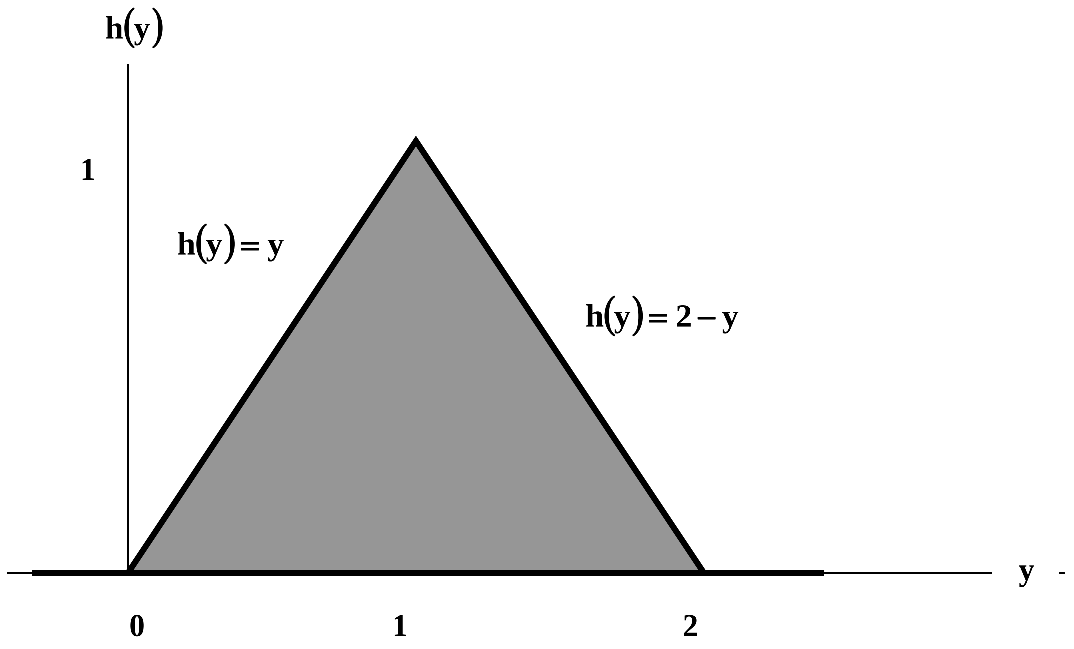
$$g(y,z) = 1.|1| = 1$$

لقيم $z < y < z + 1$ ، $0 < z < 1$ ، وبخلاف ذلك $g(y,z) = 0$.

(II) . باستخدام دالة الكثافة المشتركة في الجزء **(I)** وبإجراء التكامل علي قيم z المختلفة لقيم y التي تحقق $y \le 0$ ، $0 < y < 1$ ، $1 < y < 2$ وكذلك $y \ge 2$ نحصل علي

$$h(z) = \begin{cases} 0 & \text{for} \quad y \le 0 \\ \int_0^y 1dz = y & \text{for} \quad 0 < y < 1 \\ \int_{y-1}^1 1dz = 2 - y & \text{for} \quad 1 < y < 2 \\ 0 & \text{for} \quad y \ge 2 \end{cases}$$

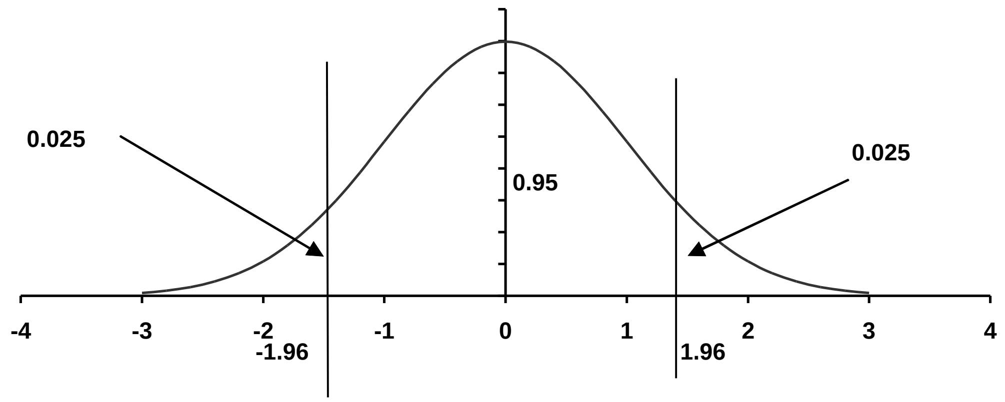
وحتى تكون دالة الكثافة متصلة فأننا نفترض أن $h(1)=1$. وبالتالي فأننا نري أن مجموع المتغيرات المعطاة له دالة الكثافة الموزعة علي المثلث كما هو موضح بالرسم



إجابة السؤال الثالث (أ):

في حالة ما يكون التباين للمجتمع معلوم ويساوي 4

عند درجة ثقة 95% اي ان $1-\alpha=0.95$ **نجد أن** $\alpha=0.05, \frac{\alpha}{2}=0.025$ **ومن الجداول نجد أن** $z_{\frac{\alpha}{2}}=1.96$ **وبالتالى فأن**



$$\bar{x}-z_{\frac{\alpha}{2}}.\frac{\sigma}{\sqrt{n}}<\mu<\bar{x}+z_{\frac{\alpha}{2}}.\frac{\sigma}{\sqrt{n}}$$

$$23.5-1.96\times\frac{2}{\sqrt{16}}<\mu<23.5+1.96\times\frac{2}{\sqrt{16}}$$

$$22.52<\mu<24.48$$

أي أن فترة الثقة المناظرة لمستوى ثقة **95%** هى $(22.52,24.48)$.

في حالة ما يكون التباين للمجتمع مجهول وتباين العينة معلوم ويساوي 5

عند درجة ثقة 99% اي ان $1-\alpha=0.99$ **نجد أن** $\alpha=0.05, \frac{\alpha}{2}=0.005$ **ومن الجداول نجد أن** $t_{(0.005,15)}=2.974$

$$\bar{x}-t_{(n-1),\frac{\alpha}{2}}\cdot\frac{s}{\sqrt{n}}<\mu<\bar{x}+t_{(n-1),\frac{\alpha}{2}}\cdot\frac{s}{\sqrt{n}}$$

$$23.5-2.974\times\frac{\sqrt{5}}{\sqrt{16}}<\mu<23.5+2.974\times\frac{\sqrt{5}}{\sqrt{16}}$$

$$22.94<\mu<24.06$$

أي أن فترة الثقة المناظرة لمستوى ثقة 99% هى $(22.94,24.06)$.

إجابة السؤال الثالث (ب):

خصائص التقدير الجيد

أ . التقدير غير المتميز : Unbiased estimator

نعتبر مجتمع معين أحد معالمه " بارامتر " هو θ هذا المجتمع أخذنا عينة عشوائية واحدة حجمها n ومنها أمكن تعين التقدير الإحصائي $\hat{\theta}$ لتقدير للبارامتر θ التقدير $\hat{\theta}$ ماهو الا متغير عشوائى يتغير من عينة إلى أخرى علما بأن حجم العينة n ثابت .يقال أن التقدير الإحصائي $\hat{\theta}$ بانه تقدير غير متحيز للبارمتر θ إذا كان $E\,\hat{\theta}=\theta$

د . الأتساق Consistency

الخاصية الثانية التى يجب أن تتوافر حتى نقول أن التقدير جيدا هو الأتساق .

نفرض أن لدينا مجتمع حجمه N وسحبنا منه عينة عشوائية حجمها n ونفرض أن معلمة المجتمع المجهولة هى θ وأن التقدير من العينة هو $\hat{\theta}$ وبالتالى يقال أن : التقدير $\hat{\theta}$ تقديرا متسقا للبارمتر θ إذا تقارب هذا التقدير للبارمتر θ عن طريق الاحتمال أى أن

$$\lim_{n\to\infty} P\left(\left|\theta-\hat{\theta}\right|\geq c\right)\to 1$$

حيث c مقدار اختيارى موجب $c>0$.

وا لتعرف السابق مكافىء إلى أن يكون يسمى التقدير $\hat{\theta}$ تقديرا متسقا للبارمتر θ إذا كان :

1 . $\hat{\theta}$ تقديرا غير متحيزا للبارمتر θ

2 . $$\lim_{n\to\infty}\sigma^2\left(\hat{\theta}\right)\to 0$$

هـ الكفاية : Sufficiency

التقدير $\hat{\theta}$ تقديرا كاف للبارمتر θ لأى مجتمع إذا أمكن التعبير عن دالة الكثافة الاحتمالية المشتركة لمفردات العينة كحاصل ضرب دالتين احدهما تعتمد فقط على البارمتر

θ والتقدير $\hat{\theta}$ والآخرى لاتعتمد على البارمتر θ $\mathbf{0}$ أى أن

$$f\left(x_1, x_2, \ldots, x_n; \theta\right) = g\left(\theta, \hat{\theta}\right) h\left(x_1, x_2, \ldots, x; \hat{\theta}\right)$$

دالة الكثافة الاحتمالية لمجتمع يتبع التوزيع الآسي هى $f(x) = \theta . e^{-\theta . x}, 0 < x < \infty$:

من هذه المعادلة يتضح أن دالة الاحتمال $f(x)$ تعتمد على معلمة واحدة وهى θ وبذلك سوف نكون معادلة واحدة وذلك بمساواة العزم الأول المحسوب من دالة الكثافة الاحتمالية بالعزم الأول المحسوب من العينة أى أن

$$\mu_1 = EX = \int_{-\infty}^{\infty} xf(x)\,dx =$$

$$= \int x\theta . f(x) e^{-\theta . x} dx = \frac{1}{\theta}$$

والعزم الأول المحسوب من العينة هو :

$$m_1 = \bar{x} = \frac{\sum x}{n}$$

و بمساواة المعادلتين ينتج أن

$$\frac{1}{\theta} = \frac{\sum x}{n}$$

أي أن

$$\hat{\theta} = \frac{n}{\sum x} = \frac{4}{60.1} = 0.067$$

وبذلك تكون دالة التوزيع المقدرة كالآتى :

$$f(x) = 0.067 e^{-0.067 x}, \qquad 0 < x < \infty$$

إجابة السؤال الرابع (أ):

هذا المجتمع له الدالة الاحتمالية

X	1	3	5	7	9
$f(X)$	1/5	1/5	1/5	1/5	1/5

يتضح أن هذا التوزيع ليس معتدلا بل يطلق عليه التوزيع المنتظم ويمكن حساب المتوسط والتباين ونجد أن

$\sigma^2 = 8; \mu = 5$

الآن نريد سحب عينة من حجم 2 ثم نوجد التوزيع العينى على أن يكون السحب بإرجاع فى هذه الحالة نجد أن عدد العينات الممكنة هى $5^2 = 25$ وهى كالتالى :

$$(1,1), (1,3), (1,5), (1,7), (1,9)$$
$$(3,1), (3,3), (3,5), (3,7), (3,9)$$
$$(5,1), (5,3), (5,5), (5,7), (5,9)$$
$$(7,1), (7,3), (7,5), (7,7), (7,9)$$
$$(9,1), (9,3), (9,5), (9,7), (9,9)$$

احتمال سحب عينة من هذه العينات هى $1/25$ ويكون التوزيع العينى هو :

$\overline{X}$	1	2	3	4	5	6	7	8	9
$P(\overline{X})$	$1/25$	$2/25$	$3/25$	$4/25$	$5/25$	$4/25$	$3/25$	$2/25$	$1/25$

ويرسم هذا التوزيع نجد أنه يأخذ تقريبا شكل التوزيع المعتدل

$$E\overline{X} = 1\times\frac{1}{25}+2\times\frac{2}{25}+3\times\frac{3}{25}+4\times\frac{4}{25}+5\times\frac{5}{25}+6\times\frac{4}{25}+7\times\frac{3}{25}+8\times\frac{2}{25}+9\times\frac{1}{25}=5$$

$$\sigma^2{}_{\overline{X}} = 1^2\times\frac{1}{25}+2^2\times\frac{2}{25}+3^2\times\frac{3}{25}+4^2\times\frac{4}{25}+5^2\times\frac{5}{25}+6^2\times\frac{4}{25}+7^2\times\frac{3}{25}+8^2\times\frac{2}{25}+9^2\times\frac{1}{25}-5^2$$
$$=29-25=4$$

ونجد أن :

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}, \ E\overline{X} = \mu$$

إجابة السؤال الرابع (ب):

نسبة المشاهدين لهذا البرنامج فى العينة $r=\frac{190}{250}=0.76$ حجم العينة **n=250**

وبالتالى فأن :

$$\sigma(r)=\sqrt{\frac{p(1-p)}{n}}$$

وحيث أن **p** مجهولة لذلك نستخدم النسبة **r** بدلا من **p**

$$\sigma(r)=\sqrt{\frac{r(1-r)}{n}}=\sqrt{\frac{0.76\times(1-0.76)}{250}}=0.027$$

وبالتالى فأن

$$P[0.76-1.96\times0.027\le p\le 0.76+1.96\times0.027]=0.95$$

أى أن

$$P(0.707\le p\le 0.813)=0.95$$

أى أننا نتوقع أن p تقع بين 0.707 & 0.813 وذلك بمستوى ثقة 95% .