Exam 2 PREP

Evaluate but only reduce if it is requested.

1. Define G(x, y) = x^2y + x, 1 < x < 5, 3 < y < 4 (elsewhere zero).

a. Determine the integral $\int\int dx\, dy\, g(x, y)$ over the plane. Reduce to a number.

b. From (a) define the probability density f(x,y) proportional to g.

c. From (b) determine the marginal density of r.v. X.

d. From (c) determine E X.

e. From (c) determine

E X^2

Var X

f. From (c) determine the probability density of r.v. $Y = X^2$ using the method

pd of Y = h(X) is $f_Y(y) = f_X(x) \,/\, |\, h'(x)\, |$

for 1:1 differentiable function y = h(x) with derivative h' ≠ 0 with probability 1.

g. From (f) determine E Y = E X^2 and compare with your calculation done in (e).

h. E(X Y)

2. Define cumulative distribution F(x) = 1 - x^{-3}, $x > 1$ (zero elsewhere).

a. Determine probability density f from cumulative F.

b. Determine E X.

3. For random variables X, Y suppose

E X = 3	E Y = 7
sd X = 1	sd Y = 4

a. E(2 X - Y + 3)

b. If X, Y are independent, Var(2 X - Y + 3)

c. E(X Y)

4. Random variables X, Y have

E X = 3	E Y = 7

sd X = 1 sd Y = 4
E (XY) = 20

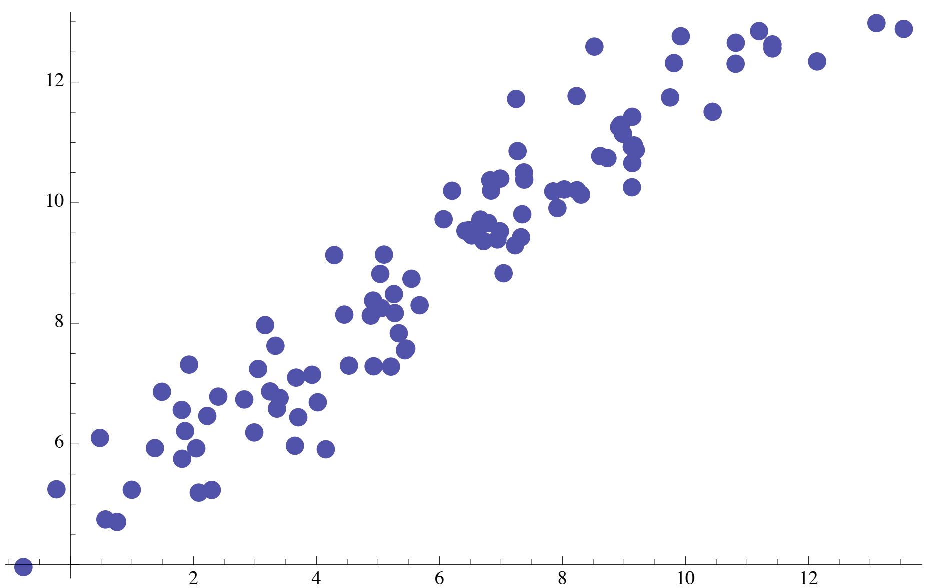
a. Covariance Cov(X, Y) = E(XY) - (E X)(E Y)

b. Correlation R = $\frac{\text{Cov}(X, Y)}{\text{sd } X \text{ sd } Y}$ (must be a number in [-1 and 1])

c. The regression line of y on x is the line passing through the point (E X, E Y) with slope R (sd Y) / (sd X). Plot this line by

laying off sd X to the right of (E X, E Y)
going up R sd Y from EY
making the line thru (E X, E Y) and (E X + sd X, E Y + R sd Y)

d. If (X, Y) are jointly normal distributed the density has elliptical contours around the regression line. Here is a plot of samples from such a joint distribution. By eye, draw in the regression line



roughly locate (E X, E Y) by eye
draw in sd X, sd Y by intervals containing ~ 68% of points (in x then y)

e. What does the plot of E(Y | X = x) vs x look like in (d)?

5. A random sample of n = 7 from a NORMAL population finds
sample mean = 111.69
sample sd s = 84.77

a. Estimate of population mean μ is

b. Estimate of population sd σ is

c. Estimate of sd of sample mean $\bar{\boldsymbol{x}}$ is

d. t-score for 95% CI for μ is

e. 95% CI for μ is

f. P(μ in 95% CI) (exactly in this case) =

g. Why can we not use this method for sample size n = 1?

h. If $\frac{\overline{x}-\mu}{s/\sqrt{n}}$=1.647 and we change location and scale to (new) x = 3 x - 4 what is the (new) value of $\frac{\overline{x}-\mu}{s/\sqrt{n}}$?

6. Bootstrap CI for population mean μ typically involves thousands of bootstrap samples X_1^*, ... , X_n^* (equal pr with repl from our sample X_1, ... , X_n). For illustration only, we will illustrate bootstrap using 20 bootstrap sample means. They are plotted on the line below as asterisks. The sample mean is a large plus sign. By eye, draw in the bootstrap 95% CI for μ.

*   *   *   **   + *   ****   * * **   ** ** * *

7. Joint pmf is given below.

x \ y	4	8
5	.3	.4
7	.2	.1

a. Give pmf (marginal density) for r.v. X.

b. E X from (a).

c. E XY

d. E(Y | X = 1)

e. Var(Y | X = 1)

f. E Var(Y | X) + Var E(Y | X) = (something simple)

g. Which of the two parts of (f) is "between" component?

8. Calculate sample sd s for data {6, 6, 14, 10}.