FIRST MIDTERM ECON 7800 FALL 2003

ECONOMICS DEPARTMENT, UNIVERSITY OF UTAH

Problem 1. *3 points Draw a Venn Diagram which shows the validity of de Morgan's laws:* $(A \cup B)' = A' \cap B'$ *and* $(A \cap B)' = A' \cup B'$. *If done right, the same Venn diagram can be used for both proofs.*

Answer. There is a proof in [HT83, p. 12]. Draw A and B inside a box which represents U, and shade A' from the left (blue) and B' from the right (yellow), so that $A' \cap B'$ is cross shaded (green); then one can see these laws. □

Problem 2. *2 points (Bonferroni inequality) Let* A *and* B *be two events. Writing* $\Pr[A] = 1 - \alpha$ *and* $\Pr[B] = 1 - \beta$, *show that* $\Pr[A \cap B] \geq 1 - (\alpha + \beta)$. *You are allowed to use that* $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$ *(Problem* **??***), and that all probabilities are* ≤ 1.

Date of exam Thursday, September 18, 2003, 12:25–1:45 pm in BUC 207.

Answer.

$$\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B] \le 1 \tag{1}$$

$$\Pr[A] + \Pr[B] \le 1 + \Pr[A \cap B] \tag{2}$$

$$\Pr[A] + \Pr[B] - 1 \le \Pr[A \cap B] \tag{3}$$

$$1 - \alpha + 1 - \beta - 1 = 1 - \alpha - \beta \le \Pr[A \cap B] \tag{4}$$

□

Problem 3. *2 points A and B are arbitrary events. Prove that the probability of B can be written as:*

$$\Pr[B] = \Pr[B|A]\Pr[A] + \Pr[B|A']\Pr[A'] \tag{5}$$

Answer. $B = B \cap U = B \cap (A \cup A') = (B \cap A) \cup (B \cap A')$ and this union is disjoint, i.e., $(B \cap A) \cap (B \cap A') = B \cap (A \cap A') = B \cap \emptyset = \emptyset$. Therefore $\Pr[B] = \Pr[B \cap A] + \Pr[B \cap A']$. Now apply definition of conditional probability to get $\Pr[B \cap A] = \Pr[B|A]\Pr[A]$ and $\Pr[B \cap A'] = \Pr[B|A']\Pr[A']$. □

Problem 4. *The first three questions here are discussed in* [Lar82, example 2.6.3 on p. 62]*: There is an urn with 4 white and 8 black balls. You take two balls out without replacement.*

- **a.** *1 point What is the probability that the first ball is white?*

Answer. 1/3 □

- **b.** *1 point What is the probability that both balls are white?*

Answer. It is $\Pr[\text{second ball white}|\text{first ball white}]\Pr[\text{first ball white}] = \frac{3}{3+8}\frac{4}{4+8} = \frac{1}{11}$. □

- **c.** *1 point What is the probability that the second ball is white?*

Answer. It is $\Pr[\text{first ball white and second ball white}]+\Pr[\text{first ball black and second ball white}] =$

$$= \frac{3}{3+8}\frac{4}{4+8} + \frac{4}{7+4}\frac{8}{8+4} = \frac{1}{3}. \tag{6}$$

This is the same as the probability that the first ball is white. The probabilities are not dependent on the order in which one takes the balls out. This property is called "exchangeability." One can see it also in this way: Assume you number the balls at random, from 1 to 12. Then the probability for a white ball to have the number 2 assigned to it is obviously $\frac{1}{3}$. □

- **d.** *1 point What is the probability that both of them are black?*

Answer. $\frac{8}{12}\frac{7}{11} = \frac{2}{3}\frac{7}{11} = \frac{14}{33}$ (or $\frac{56}{132}$). □

- **e.** *1 point What is the probability that both of them have the same color?*

Answer. The sum of the two above, $\frac{14}{33} + \frac{1}{11} = \frac{17}{33}$ (or $\frac{68}{132}$). □

Problem 5. *2 points A friend tosses two coins. You ask: "did one of them land heads?" Your friend answers, "yes." What's the probability that the other also landed heads?*

Answer. $U = \{HH, HT, TH, TT\}$; Probability is $\frac{1}{4}/\frac{3}{4} = \frac{1}{3}$. □

Problem 6. *AIDS diagnostic tests are usually over 99.9% accurate on those who do not have AIDS (i.e., only 0.1% false positives) and 100% accurate on those who have AIDS (i.e., no false negatives at all). (A test is called positive if it indicates that the subject has AIDS.)*

• **a.** *3 points Assuming that 0.5% of the population actually have AIDS, compute the probability that a particular individual has AIDS, given that he or she has tested positive.*

Answer. A is the event that he or she has AIDS, and T the event that the test is positive.

$$\Pr[A|T] = \frac{\Pr[T|A]\Pr[A]}{\Pr[T|A]\Pr[A] + \Pr[T|A']\Pr[A']} = \frac{1 \cdot 0.005}{1 \cdot 0.005 + 0.001 \cdot 0.995} =$$
$$= \frac{100 \cdot 0.5}{100 \cdot 0.5 + 0.1 \cdot 99.5} = \frac{1000 \cdot 5}{1000 \cdot 5 + 1 \cdot 995} = \frac{5000}{5995} = \frac{1000}{1199} = 0.834028$$

Even after testing positive there is still a 16.6% chance that this person does not have AIDS. □

• **b.** *1 point If one is young, healthy and not in one of the risk groups, then the chances of having AIDS are not 0.5% but 0.1% (this is the proportion of the applicants to the military who have AIDS). Re-compute the probability with this alternative number.*

Answer.

$$\frac{1 \cdot 0.001}{1 \cdot 0.001 + 0.001 \cdot 0.999} = \frac{100 \cdot 0.1}{100 \cdot 0.1 + 0.1 \cdot 99.9} = \frac{1000 \cdot 1}{1000 \cdot 1 + 1 \cdot 999} = \frac{1000}{1000 + 999} = \frac{1000}{1999} = 0.50025.$$

□

Problem 7. *2 points A and B are two independent events with $\Pr[A] = \frac{1}{3}$ and $\Pr[B] = \frac{1}{4}$. Compute $\Pr[A \cup B]$.*

Answer. $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B] = \Pr[A] + \Pr[B] - \Pr[A]\Pr[B] = \frac{1}{3} + \frac{1}{4} - \frac{1}{12} = \frac{1}{2}$. □

Problem 8. *2 points Flip a coin two times independently and define the following three events:*

$$\begin{aligned} A &= \text{Head in first flip} \\ B &= \text{Head in second flip} \\ C &= \text{Same face in both flips.} \end{aligned} \tag{7}$$

Are these three events pairwise independent? Are they mutually independent?

Answer. $U = \left\{ \begin{smallmatrix} HH & HT \\ TH & TT \end{smallmatrix} \right\}$. $A = \{HH, HT\}$, $B = \{HH, TH\}$, $C = \{HH, TT\}$. $\Pr[A] = \frac{1}{2}$, $\Pr[B] = \frac{1}{2}$, $\Pr[C] = \frac{1}{2}$. They are pairwise independent, but $\Pr[A \cap B \cap C] = \Pr[\{HH\}] = \frac{1}{4} \neq \Pr[A]\Pr[B]\Pr[C]$, therefore the events cannot be mutually independent. □

Problem 9. *4 points Assume the random variable z has the exponential distribution with $\lambda = 1$, i.e., its density function is $f_z(z) = \exp(-z)$ for $z \geq 0$ and 0 for $z < 0$. Define $u = \sqrt{z}$. Compute the density function of u.*

Answer. (1) $A = \{u\colon u \geq 0\}$ since $\sqrt{\ }$ always denotes the nonnegative square root; (2) Express old variable in terms of new: $z = u^2$, this is one-to-one on A (but not one-to-one on all of $\mathbb{R}$); (3) then the derivative is $2u$, which is nonnegative as well, no absolute values are necessary; (4) multiplying gives the density of u: $f_u(u) = 2u\exp(-u^2)$ if $u \geq 0$ and 0 elsewhere. □

Problem 10. *You perform a Bernoulli experiment, i.e., an experiment which can only have two outcomes, success s and failure f. The probability of success is p.*

- **a.** *3 points You make 4 independent trials. Show that the probability that the first trial is successful, given that the total number of successes in the 4 trials is 3, is 3/4.*

Answer. Let $B = \{sfff, sffs, sfsf, sfss, ssff, ssfs, sssf, ssss\}$ be the event that the first trial is successful, and let $\{x{=}3\} = \{fsss, sfss, ssfs, sssf\}$ be the event that there are 3 successes, it has $\binom{4}{3} = 4$ elements. Then

$$\Pr[B|x{=}3] = \frac{\Pr[B \cap \{x{=}3\}]}{\Pr[x{=}3]} \tag{8}$$

Now $B \cap \{x{=}3\} = \{sfss, ssfs, sssf\}$, which has 3 elements. Therefore we get

$$\Pr[B|x{=}3] = \frac{3 \cdot p^3(1-p)}{4 \cdot p^3(1-p)} = \frac{3}{4}. \tag{9}$$

□

- **b.** *2 points Discuss this result.*

Answer. It is significant that this probability is independent of p. I.e., once we know how many successes there were in the 4 trials, knowing the true p does not help us computing the probability of the event. From this also follows that the outcome of the event has no information about p. The value 3/4 is the same as the unconditional probability if $p = 3/4$. I.e., whether we know that the true frequency, the one that holds in the long run, is 3/4, or whether we know that the actual frequency in this sample is 3/4, both will lead us to the same predictions regarding the first throw. But not all conditional probabilities are equal to their unconditional counterparts: the conditional probability to get 3 successes in the first 4 trials is 1, but the unconditional probability is of course not 1. □

Problem 11. *4 points Assume your data show that counties with high rates of unemployment also have high rates of heart attacks. Can one conclude from this that the unemployed have a higher risk of heart attack? Discuss, besides the "ecological fallacy," also other objections which one might make against such a conclusion.*

Answer. Ecological fallacy says that such a conclusion is only legitimate if one has individual data. Perhaps a rise in unemployment is associated with increased pressure and increased workloads among the employed, therefore it is the employed, not the unemployed, who get the heart attacks. Even if one has individual data one can still raise the following objection: perhaps unemployment and heart attacks are both consequences of a third variable (both unemployment and heart attacks depend on age or education, or freezing weather in a farming community causes unemployment for workers and heart attacks for the elderly).

□

Problem 12. [CT91, example 2.1.2 on pp. 14/15]: *The experiment has four possible outcomes; outcome* $x=a$ *occurs with probability* 1/2, $x=b$ *with probability* 1/4, $x=c$ *with probability* 1/8, *and* $x=d$ *with probability* 1/8.

• **a.** *2 points For this part you will need*

$$\frac{\mathrm{H}[\mathcal{F}]}{\text{bits}} = \sum_{k=1}^{n} p_k \log_2 \frac{1}{p_k} \tag{10}$$

The entropy of this experiment (in bits*) is one of the following three numbers: 11/8, 7/4, 2. Which is it?*

• **b.** *2 points Suppose we wish to determine the outcome of this experiment with the minimum number of questions. An efficient first question is "Is* $x=a$*?" This splits the probability in half. If the answer to the first question is no, then the second question can be "Is* $x=b$*?" The third question, if it is necessary, can then be: "Is* $x=c$*?" Compute the expected number of binary questions required.*

• **c.** *2 points Show that the entropy gained by each question is 1 bit.*

• **d.** *3 points Assume we know about the first outcome that* $x \neq a$*. What is the entropy of the remaining experiment (i.e., under the conditional probability)?*

• **e.** *5 points Show in this example that the composition law for entropy holds.*

Problem 13. *2 points An exponential random variable t with parameter $\lambda > 0$ has the density $f_t(t) = \lambda e^{-\lambda t}$ for $t \geq 0$, and 0 for $t < 0$. Use this density to compute the expected value of t.*

Answer. $\mathrm{E}[t] = \int_0^\infty \lambda t e^{-\lambda t} dt = \int_0^\infty uv' dt = uv\Big|_0^\infty - \int_0^\infty u'v dt$, where $\begin{smallmatrix} u=t & v'=\lambda e^{-\lambda t} \\ u'=1 & v=-e^{-\lambda t} \end{smallmatrix}$. One can also use the more abbreviated notation $= \int_0^\infty u\, dv = uv\Big|_0^\infty - \int_0^\infty v\, du$, where $\begin{smallmatrix} u=t & dv'=\lambda e^{-\lambda t}\, dt \\ du'=dt & v=-e^{-\lambda t} \end{smallmatrix}$. Either way one obtains $\mathrm{E}[t] = -te^{-\lambda t}\Big|_0^\infty + \int_0^\infty e^{-\lambda t} dt = 0 - \frac{1}{\lambda} e^{-\lambda t}|_0^\infty = \frac{1}{\lambda}$. □

Problem 14. *Let x be uniformly distributed in the interval $[a, b]$, i.e., the density function of x is a constant for $a \leq x \leq b$, and zero otherwise.*

- **a.** *1 point What is the value of this constant?*

Answer. It is $\frac{1}{b-a}$ □

- **b.** *2 points Compute $\mathrm{E}[x]$*

Answer. $\mathrm{E}[x] = \int_a^b \frac{x}{b-a}\, dx = \frac{1}{b-a} \frac{b^2-a^2}{2} = \frac{a+b}{2}$ since $b^2 - a^2 = (b+a)(b-a)$. □

- **c.** *2 points Show that $\mathrm{E}[x^2] = \frac{a^2+ab+b^2}{3}$.*

Answer. $\mathrm{E}[x^2] = \int_a^b \frac{x^2}{b-a}\, dx = \frac{1}{b-a} \frac{b^3-a^3}{3}$. Now use the identity $b^3 - a^3 = (b-a)(b^2 + ab + a^2)$ (check it by multiplying out). □

• **d.** *2 points Show that* $\operatorname{var}[x] = \frac{(b-a)^2}{12}$.

Answer. $\operatorname{var}[x] = \operatorname{E}[x^2] - (\operatorname{E}[x])^2 = \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4} = \frac{4a^2+4ab+4b^2}{12} - \frac{3a^2+6ab+3b^2}{12} = \frac{(b-a)^2}{12}$. □

Problem 15. *3 points Explain verbally clearly what the law of large numbers means, what the Central Limit Theorem means, and what their difference is.*

Problem 16. *3 points For two random variables* x, y, *their* covariance *is defined as*

$$\operatorname{cov}[x, y] = \operatorname{E}\Big[\big(x - \operatorname{E}[x]\big)\big(y - \operatorname{E}[y]\big)\Big]. \tag{11}$$

Using this definition, prove the following formula:

$$\operatorname{cov}[x, y] = \operatorname{E}[xy] - \operatorname{E}[x]\operatorname{E}[y]. \tag{12}$$

Write it down carefully, you will lose points for unbalanced or missing parantheses and brackets.

Answer. Here it is side by side with and without the notation $\operatorname{E}[x] = \mu$ and $\operatorname{E}[y] = \nu$:

$$\begin{aligned} \operatorname{cov}[x, y] &= \operatorname{E}\Big[(x - \operatorname{E}[x])(y - \operatorname{E}[y])\Big] \\ &= \operatorname{E}\Big[xy - x\operatorname{E}[y] - \operatorname{E}[x]y + \operatorname{E}[x]\operatorname{E}[y]\Big] \\ &= \operatorname{E}[xy] - \operatorname{E}[x]\operatorname{E}[y] - \operatorname{E}[x]\operatorname{E}[y] + \operatorname{E}[x]\operatorname{E}[y] \end{aligned} \qquad \begin{aligned} \operatorname{cov}[x, y] &= \operatorname{E}[(x - \mu)(y - \nu)] \\ &= \operatorname{E}[xy - x\nu - \mu y + \mu\nu] \\ &= \operatorname{E}[xy] - \mu\nu - \mu\nu + \mu\nu \end{aligned} \tag{13}$$

□

Problem 17. *2 points The conditional density is the joint divided by the marginal:*

$$f_{y|x}(y, x) = \frac{f_{x,y}(x, y)}{f_x(x)}. \tag{14}$$

Show that this density integrates out to 1.

Answer. The conditional is a density in y with x as parameter. Therefore its integral with respect to y must be $= 1$. Indeed,

$$\int_{y=-\infty}^{+\infty} f_{y|x=x}(y, x)\, dy = \frac{\int_{y=-\infty}^{+\infty} f_{x,y}(x, y)\, dy}{f_x(x)} = \frac{f_x(x)}{f_x(x)} = 1 \tag{15}$$

because of the formula for the marginal:

$$f_x(x) = \int_{y=-\infty}^{+\infty} f_{x,y}(x, y)\, dy \tag{16}$$

You see that formula (14) divides the joint density exactly by the right number which makes the integral equal to 1. □

Problem 18. *3 points What is the best predictor of a random variable y by a constant a, if the loss function is the "mean squared error" (MSE)* $\mathrm{E}[(y - a)^2]$*?*

Answer. Write $\mathrm{E}[y] = \mu$; then

$$\begin{aligned} (y-a)^2 &= \big((y-\mu) - (a-\mu)\big)^2 \\ &= (y-\mu)^2 - 2(y-\mu)(a-\mu) + (a-\mu)^2; \end{aligned} \tag{17}$$

$$\text{therefore} \quad \mathrm{E}[(y-a)^2] = \mathrm{E}[(y-\mu)^2] - 0 + (a-\mu)^2$$

This is minimized by $a = \mu$. □

Problem 19. *Assume the vector $\boldsymbol{x} = [x_1, \dots x_j]^\top$ and the scalar y are jointly distributed random variables, and assume conditional means exist. Define $\varepsilon = y - \mathrm{E}[y|\boldsymbol{x}]$.*

For this problem you may use (1) the theorem of iterated expectations $\mathrm{E}\big[\mathrm{E}[y|\boldsymbol{x}]\big] = \mathrm{E}[y]$, (2) the additivity $\mathrm{E}[g(y)+h(y)|\boldsymbol{x}] = \mathrm{E}[g(y)|\boldsymbol{x}] + \mathrm{E}[h(y)|\boldsymbol{x}]$, and (3) the fact that $\mathrm{E}[g(\boldsymbol{x})h(y)|\boldsymbol{x}] = g(\boldsymbol{x})\mathrm{E}[h(y)|\boldsymbol{x}]$.

- **a.** *5 points Demonstrate the following identities:*

$$\mathrm{E}[\varepsilon|\boldsymbol{x}] = 0 \tag{18}$$

$$\mathrm{E}[\varepsilon] = 0 \tag{19}$$

$$\mathrm{E}[x_i\varepsilon|\boldsymbol{x}] = 0 \quad \textit{for all } i,\ 1 \le i \le j \tag{20}$$

$$\mathrm{E}[x_i\varepsilon] = 0 \quad \textit{for all } i,\ 1 \le i \le j \tag{21}$$

$$\operatorname{cov}[x_i, \varepsilon] = 0 \quad \textit{for all } i,\ 1 \le i \le j. \tag{22}$$

Interpretation of (22): ε is the error in the best prediction of y based on $\boldsymbol{x}$. If this error were correlated with one of the components x_i, then this correlation could be used to construct a better prediction of y.

Answer. (18): $\mathrm{E}[\varepsilon|\boldsymbol{x}] = \mathrm{E}[y|\boldsymbol{x}] - \mathrm{E}\Big[\mathrm{E}[y|\boldsymbol{x}]|\boldsymbol{x}\Big] = 0$ since $\mathrm{E}[y|\boldsymbol{x}]$ is a function of $\boldsymbol{x}$ and therefore equal to its own expectation conditionally on $\boldsymbol{x}$. (This is *not* the law of iterated expectations but the law that the expected value of a constant is a constant.)

(19) follows from (18) (i.e., (18) is stronger than (19)): if an expectation is zero conditionally on every possible outcome of $\boldsymbol{x}$ then it is zero altogether. In formulas, $\mathrm{E}[\varepsilon] = \mathrm{E}\Big[\mathrm{E}[\varepsilon|\boldsymbol{x}]\Big] = \mathrm{E}[0] = 0$. It is also easy to show it in one swoop, without using (18): $\mathrm{E}[\varepsilon] = \mathrm{E}[y - \mathrm{E}[y|\boldsymbol{x}]] = 0$. Either way you need the law of iterated expectations for this.

(20): $\mathrm{E}[x_i\varepsilon|\boldsymbol{x}] = x_i\mathrm{E}[\varepsilon|\boldsymbol{x}] = 0$.

(21): $\mathrm{E}[x_i\varepsilon] = \mathrm{E}\Big[\mathrm{E}[x_i\varepsilon|\boldsymbol{x}]\Big] = \mathrm{E}[0] = 0$; or in one swoop: $\mathrm{E}[x_i\varepsilon] = \mathrm{E}\Big[x_iy - x_i\mathrm{E}[y|\boldsymbol{x}]\Big] = \mathrm{E}\Big[x_iy - \mathrm{E}[x_iy|\boldsymbol{x}]\Big] = \mathrm{E}[x_iy] - \mathrm{E}[x_iy] = 0$. The following "proof" is not correct: $\mathrm{E}[x_i\varepsilon] = \mathrm{E}[x_i]\,\mathrm{E}[\varepsilon] = \mathrm{E}[x_i] \cdot 0 = 0$. x_i and ε are generally not independent, therefore the multiplication rule $\mathrm{E}[x_i\varepsilon] = \mathrm{E}[x_i]\,\mathrm{E}[\varepsilon]$ cannot be used. Of course, the following "proof" does not work either: $\mathrm{E}[x_i\varepsilon] = x_i\,\mathrm{E}[\varepsilon] = x_i \cdot 0 = 0$. x_i is a random variable and $\mathrm{E}[x_i\varepsilon]$ is a constant; therefore $\mathrm{E}[x_i\varepsilon] = x_i\,\mathrm{E}[\varepsilon]$ cannot hold.

(22): $\operatorname{cov}[x_i, \varepsilon] = \mathrm{E}[x_i\varepsilon] - \mathrm{E}[x_i]\,\mathrm{E}[\varepsilon] = 0 - \mathrm{E}[x_i] \cdot 0 = 0$. □

• **b.** *2 points If $\boldsymbol{x}$ and y are jointly normal, show that $\boldsymbol{x}$ and ε are independent, and that the variance of ε does not depend on $\boldsymbol{x}$. (This is why one can consider it an error term.)*

Answer. If $\boldsymbol{x}$ and y are jointly normal, then $\boldsymbol{x}$ and ε are jointly normal as well, and independence follows from the fact that their covariance is zero. The variance is constant because in the Normal case, the conditional variance is constant, i.e., $\mathrm{E}[\varepsilon^2] = \mathrm{E}\Big[\mathrm{E}[\varepsilon^2|\boldsymbol{x}]\Big] =$ constant (does not depend on $\boldsymbol{x}$). □

Problem 20. *The evaluation of two intelligence tests, one at the beginning of the semester, one at the end, gives the following disturbing outcome: While the underlying intelligence during the first test was $z \sim N(100, 20)$, it changed between the first and second test due to the learning experience at the university. If w is the intelligence of each student at the second test, it is connected to his intelligence z at the first test by the formula $w = 0.5z + 50$, i.e., those students with intelligence below 100 gained, but those students with intelligence above 100 lost. (The errors of both intelligence tests are normally distributed with expected value zero, and the variance of the first intelligence test was 5, and that of the second test, which had more questions, was 4. As usual, the errors are independent of each other and of the actual intelligence.)*

• **a.** *3 points If x and y are the outcomes of the first and second intelligence test, compute $\mathrm{E}[x]$, $\mathrm{E}[y]$, $\mathrm{var}[x]$, $\mathrm{var}[y]$, and the correlation coefficient $\rho = \mathrm{corr}[x, y]$. Figure 1 shows an equi-density line of their joint distribution; 95% of the probability mass of the test results are inside this ellipse. Draw the line $w = 0.5z + 50$ into Figure 1.*

Answer. We know $z \sim N(100, 20)$; $w = 0.5z + 50$; $x = z + \varepsilon$; $\varepsilon \sim N(0, 4)$; $y = w + \delta$; $\delta \sim N(0, 5)$; therefore $\mathrm{E}[x] = 100$; $\mathrm{E}[y] = 100$; $\mathrm{var}[x] = 20 + 5 = 25$; $\mathrm{var}[y] = 5 + 4 = 9$; $\mathrm{cov}[x, y] = 10$; $\mathrm{corr}[x, y] = 10/15 = 2/3$. In matrix notation

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim N\Big[\begin{bmatrix} 100 \\ 100 \end{bmatrix}, \begin{bmatrix} 25 & 10 \\ 10 & 9 \end{bmatrix}\Big] \tag{23}$$

The line $y = 50 + 0.5x$ goes through the points $(80, 90)$ and $(120, 110)$. □

• **b.** *4 points Here is the general formula for the conditional mean of two jointly normal variables u and v with mean μ_u and μ_v, standard deviations σ_u and σ_v, and correlation coefficient ρ:*

$$\mathrm{E}[v|u = u] = \mu_v + \rho\frac{\sigma_v}{\sigma_u}(u - \mu_u). \tag{24}$$

Since $\rho = \frac{\mathrm{cov}[u,v]}{\sigma_u \sigma_v}$, (24) can als be written as follows:

$$\mathrm{E}[v|u = u] = \mathrm{E}[v] + \frac{\mathrm{cov}[u, v]}{\mathrm{var}[u]}(u - \mathrm{E}[u]). \tag{25}$$

Compute $\mathrm{E}[y|x{=}x]$ and $\mathrm{E}[x|y{=}y]$. The first is a linear function of x and the second a linear function of y. Draw the two lines representing these linear functions into Figure 1.

Answer.

$$\mathrm{E}[y|x{=}x] = 100 + \frac{10}{25}(x - 100) = 60 + \frac{2}{5}x \tag{26}$$

$$\mathrm{E}[x|y{=}y] = 100 + \frac{10}{9}(y - 100) = -\frac{100}{9} + \frac{10}{9}y. \tag{27}$$

The line $y = \mathrm{E}[y|x{=}x]$ goes through the points $(80, 92)$ and $(120, 108)$ at the edge of Figure 1; it intersects the ellipse where it is vertical. The line $x = \mathrm{E}[x|y{=}y]$ goes through the points $(80, 82)$ and $(120, 118)$, which are the corner points of Figure 1; it intersects the ellipse where it is horizontal. The two lines intersect in the center of the ellipse, i.e., at the point $(100, 100)$.

□

- **c.** *2 points Another researcher says that* $w = \frac{6}{10}z + 40$, $z \sim N(100, \frac{100}{6})$, $\varepsilon \sim N(0, \frac{50}{6})$, $\delta \sim N(0, 3)$. *Is this compatible with the data?*

Answer. Yes, it is compatible: $\mathrm{E}[x] = \mathrm{E}[z] + \mathrm{E}[\varepsilon] = 100$; $\mathrm{E}[y] = \mathrm{E}[w] + \mathrm{E}[\delta] = \frac{6}{10}100 + 40 = 100$; $\mathrm{var}[x] = \frac{100}{6} + \frac{50}{6} = 25$; $\mathrm{var}[y] = \left(\frac{6}{10}\right)^2 \mathrm{var}[z] + \mathrm{var}[\delta] = \frac{63}{100}\frac{100}{6} + 3 = 9$; $\mathrm{cov}[x, y] = \frac{6}{10}\mathrm{var}[z] = 10$. □

- **d.** *4 points A third researcher asserts that the IQ of the students really did not change. He says* $w = z$, $z \sim N(100, 5)$, $\varepsilon \sim N(0, 20)$, $\delta \sim N(0, 4)$. *Is this compatible with the data? Is there unambiguous evidence in the data that the IQ declined?*

Answer. This is not compatible. This scenario gets everything right except the covariance: $\mathrm{E}[x] = \mathrm{E}[z] + \mathrm{E}[\varepsilon] = 100$; $\mathrm{E}[y] = \mathrm{E}[z] + \mathrm{E}[\delta] = 100$; $\mathrm{var}[x] = 5 + 20 = 25$; $\mathrm{var}[y] = 5 + 4 = 9$; $\mathrm{cov}[x, y] = 5$. A scenario in which both tests have same underlying intelligence cannot be found. Since the two conditional expectations are on the same side of the diagonal, the hypothesis that the intelligence did not change between the two tests is not consistent with the joint distribution of x and y. The diagonal goes through the points $(82, 82)$ and $(118, 118)$, i.e., it intersects the two horizontal boundaries of Figure 1. □

Maximum number of points: 89.

References

[CT91] Thomas M. Cover and Joy A. Thomas, *Elements of information theory*, Series in Telecommunications, Wiley, New York, 1991. 8

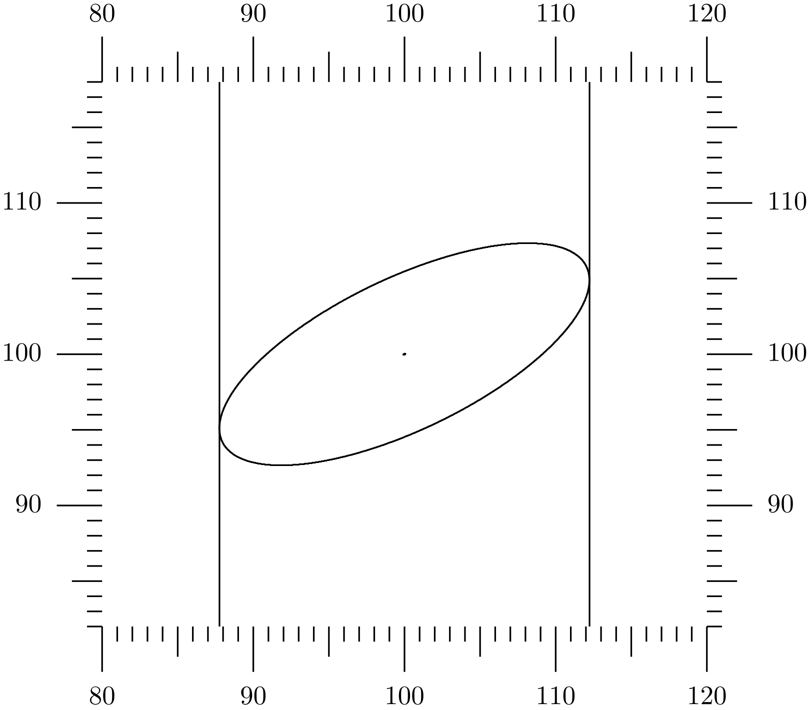
[HT83] Robert V. Hogg and Elliot A. Tanis, *Probability and statistical inference*, second ed., Macmillan, 1983. 1

[Lar82] Harold Larson, *Introduction to probability and statistical inference*, Wiley, 1982. 2

Economics Department, University of Utah, 1645 Campus Center Drive, Salt Lake City, UT 84112-9300, U.S.A

E-mail address: `ehrbar@econ.utah.edu`

URL: `http://www.econ.utah.edu/ehrbar`


80
90
100
110
120
110
100
90
110
100
90
80
90
100
110
120