Economics Prelims Probability & Statistics

Hilary Term Collection

January 2025

Instructions: You have 90 minutes to complete this closed-book, closed-notes collection. Graphing/programmable calculators are not permitted, but scientific or ordinary four-function calculators are allowed. Show your work for full credit. A table of the standard normal distribution is attached to this document.

Question:	1	2	3	4	5	6	7	Total
Points:	5	10	10	10	5	30	30	100
Score:								

☺☺☺ **Good Luck!** ☺☺☺

Part A

Answer each of the following short questions, many of which are taken *verbatim* from your review questions.

5 1. Let X be a random variable with support set $\{-1, 0, +1\}$, $p(-1) = 1/4$ and $p(0) = 1/2$. Write down the CDF of X.

Solution: By the complement rule $p(1) = 1-p(0)-p(-1) = 1-1/2-1/4 = 1/4$. Therefore,

$$F_X(x) = \begin{cases} 0, & x < -1 \\ 1/4 & -1 \leq x < 0 \\ 3/4 & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

2. Let X and Y be random variables and a, b, c be constants. Use the linearity of expectation to derive the following results

5 (a) $\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$

Solution: Let $\mu_X \equiv \mathbb{E}(X)$ and $\mu_Y \equiv \mathbb{E}(Y)$. By the definition of covariance,

$$\text{Cov}(X, Y) \equiv \mathbb{E}\left[(X - \mu_X)(Y - \mu_Y)\right].$$

Expanding and using the linearity of expectation,

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}\left[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y\right] \\ &= \mathbb{E}(XY) - \mu_X \mathbb{E}(Y) - \mu_Y \mathbb{E}(X) + \mu_X \mu_Y \\ &= \mathbb{E}(XY) - \mu_X \mu_Y - \mu_Y \mu_X + \mu_X \mu_Y \\ &= \mathbb{E}(XY) - 2\mu_X \mu_Y + \mu_X \mu_Y \\ &= \mathbb{E}(XY) - \mu_X \mu_Y. \end{aligned}$$

5 (b) $\text{Var}(a + bX + cY) = b^2\text{Var}(X) + c^2\text{Var}(Y) + 2bc\text{Cov}(X, Y)$

Solution: Let $Z \equiv a + bX + cY$ and $\mu_Z \equiv \mathbb{E}(Z)$. Define μ_X and μ_Y as in the preceding part. By the definition of variance,

$$\text{Var}(Z) \equiv \mathbb{E}\left[(Z - \mu_Z)^2\right]$$

and by the linearity of expectation

$$\mu_Z = a + b\mu_X + c\mu_Y.$$

Therefore, by the linearity of expectation and the definitions of variance and covariance, we have

$$\begin{aligned}\mathrm{Var}(Z) &= \mathbb{E}\left[\{(a+bX+cY)-(a+b\mu_X+c\mu_Y)\}^2\right]\\ &= \mathbb{E}\left[\{b(X-\mu_X)+c(Y-\mu_Y)\}^2\right]\\ &= \mathbb{E}\left[b^2(X-\mu_X)^2+c^2(Y-\mu_Y)^2+2bc(X-\mu_X)(Y-\mu_Y)\right]\\ &= b^2\mathbb{E}\left[(X-\mu_X)^2\right]+c^2\mathbb{E}\left[(Y-\mu_Y)^2\right]+2bc\mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right]\\ &= b^2\mathrm{Var}(X)+c^2\mathrm{Var}(Y)+2bc\mathrm{Cov}(X,Y)\end{aligned}$$

To answer the next question you will need to consult the normal tables appended to this document.

3. Let $X \sim \mathrm{N}(0,1)$. Calculate the following:

[2] (a) $\mathbb{P}(X=0)$

Solution: Zero: since X is continuous, $\mathbb{P}(X=x)=0$ for any realization x.

[2] (b) $\mathbb{P}(X \leq 1)$

Solution: `pnorm(1)` ≈ 0.84

[3] (c) $\mathbb{P}(-1.2 \leq X \leq 1.2)$

Solution: $\texttt{pnorm}(1.2) - \texttt{pnorm}(-1.2) \approx 0.77$

[3] (d) The value of t such that $\mathbb{P}(X > t) = 0.8$

Solution: $\texttt{qnorm}(0.2) \approx -0.84$

4. Let X be a Uniform$(0,1)$ random variable.

[5] (a) Calculate $\mathbb{E}(X)$.

Solution:

$$\mathbb{E}(X) \equiv \int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x = \int_0^1 x \cdot 1\,\mathrm{d}x = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2}$$

[5] (b) Calculate $\mathbb{E}(X^2)$

Solution:

$$\mathbb{E}(X^2) \equiv \int_{-\infty}^{+\infty} x^2 f(x)\,\mathrm{d}x = \int_0^1 x^2 \cdot 1\,\mathrm{d}x = \left.\frac{x^3}{3}\right|_0^1 = \frac{1}{3}$$

[5] 5. Ruth places one ball into her bag in secret: it is red with probability r and blue with probability $(1-r)$. You watch her place a second ball in the bag: this second ball is red. You shake the bag, reach in, and pull out a ball. The ball you draw is red. Based on this information, calculate the probability that the other ball, the one still in the bag, is *also red*. Express your answer in terms of r.

Solution: Suppose Ruth initially put a red ball in the bag. In this case, you are *certain* to draw a red ball:

$$\mathbb{P}(\text{Draw Red}|\text{Initial Ball Red}) \equiv \mathbb{P}(D_R|I_R) = 1$$

If instead she initially put a blue ball in the bag, this probability falls to 1/2.

$$\mathbb{P}(\text{Draw Red}|\text{Initial Ball Blue}) \equiv \mathbb{P}(D_R|I_B) = 1/2$$

Now: r is the prior probability that Ruth initially put a red ball in the bag, i.e. $r \equiv \mathbb{P}(I_R)$. By Bayes' Theorem,

$$\mathbb{P}(I_R|D_R) = \frac{\mathbb{P}(D_R|I_R)r}{\mathbb{P}(D_R|I_R)r + \mathbb{P}(D_R|I_B)(1-r)} = \frac{r}{r+(1-r)/2} = \frac{2r}{1+r}.$$

Part B

Answer both of these questions. They are more challenging than those from Part A, and ask you to apply what you've learned to more complicated situations.

6. Let X_1 and X_2 be independent Bernoulli random variables, each with probability of success equal to p and define the random variable Y as follows:

$$Y = \begin{cases} -1 & \text{if } X_1 = X_2 \\ 0 & \text{if } X_1 = 1 \text{ and } X_2 = 0 \\ 1 & \text{if } X_1 = 0 \text{ and } X_2 = 1 \end{cases}$$

[5] (a) Write out the joint pmf of (X_1, X_2) in a 2×2 table with values of X_1 in the rows and values of X_2 in the columns.

Solution:

		X_2	
		0	1
X_1	0	$(1-p)^2$	$p(1-p)$
	1	$p(1-p)$	p^2

5 (b) Using your answer to the preceding part, calculate the marginal pmf of Y.

Solution: Since the probabilities of mutually exclusive events sum,

$$\begin{aligned} p_Y(-1) \equiv \mathbb{P}(X_1 = X_2) &= \mathbb{P}(X_1 = 0, X_2 = 0) + \mathbb{P}(X_1 = 1, X_2 = 1) \\ &= (1-p)^2 + p^2 = (1 - 2p + p^2) + p^2 = 1 - 2p + 2p^2 \\ &= 1 - 2p(1-p) \\ p_Y(0) &= \mathbb{P}(X_1 = 1, X_2 = 0) = p(1-p) \\ p_Y(1) &= \mathbb{P}(X_1 = 0, X_2 = 1) = p(1-p). \end{aligned}$$

5 (c) Calculate $\mathbb{P}(Y = 1 | Y \neq -1)$.

Solution: Since the events $Y = -1$, $Y = 0$, and $Y = 1$ are mutually exclusive, the definition of conditional probability simplifies to:

$$\begin{aligned} \mathbb{P}(Y = 1 | Y \neq -1) &= \frac{\mathbb{P}(Y = 1 \cap Y \neq -1)}{\mathbb{P}(Y \neq -1)} = \frac{\mathbb{P}(Y = 1)}{\mathbb{P}(Y = 0) + \mathbb{P}(Y = 1)} \\ &= \frac{p(1-p)}{p(1-p) + p(1-p)} = \frac{p(1-p)}{2p(1-p)} = \frac{1}{2} \end{aligned}$$

5 (d) Frank and Ruth are playing tennis and need a fair way to decide who serves first. They decide to flip a coin, but unfortunately the only one they can find is *biased*: it comes up heads with probability $p \neq 1/2$. Based on your results from the preceding parts, suggest a way that Frank and Ruth can *simulate* a fair coin toss by tossing their biased coin *more than once*.

Solution: They should flip their unfair coin *in pairs*. If the two flips in a pair come up the same (both heads or both tails), then discard that entire pair and flip twice again from scratch (i.e. do not reuse any flip). The first time a pair of flips disagrees, the event $Y \neq -1$ has occurred. Conditional on getting a pair that disagrees, the outcome is equally likely to be (Heads, Tails) or (Tails, Heads), so the probability of "the second flip being heads" is $1/2$ regardless of the bias of the coin.

10 (e) Fun trivia fact: the method you proposed in the preceding part is often called the

von Neumann extractor, which sounds extremely bad-ass. If Frank and Ruth use this method, how many times will they need to flip their unfair coin–on average–to determine which of them will serve first? Explain how your result depends on p and why. Even if you can't calculate the expectation exactly, try to say as much as you can about it.

Solution: First things first: we know that they'll have to flip *at least twice*. They could get a disagreement on the first two flips, in which case they're done. Otherwise they'll need to flip more than twice. So the expected number of flips *must* be greater than 2. Furthermore, they'll always have to flip and *even number of times* but it could take then an *arbitrarily large number of flips*: they might get unlucky and keep getting matching results in each pair. We know that the probability of matching results in a given pair is $p_Y(-1) = 1 - 2p(1-p)$, so the probability of getting a pair that *disagrees* is $1 - p_Y(-1) = 2p(1-p)$. This probability is *smaller* when p is farther from $1/2$, since $p(1-p)$, viewed as a function of p, is a parabola that opens downward and peaks at $p = 1/2$. Taking this into account, the expected value should be *larger* when p is farther from $1/2$. In words: they'll need more flips, on average, when their coin is more biased.

To calculate the precise value of the expectation, we can use a fact that we learned in our tutorials about the Geometric(q) random variable that we discussed in our tutorials. Frank and Ruth will make *pairs of flips* until they find the first pair that disagrees ($Y \neq -1$). As long as each pair uses completely independent flips, we can view the pairs as a sequence of independent Bernoulli random variable and define a "success" to be a disagreement between a given pair. From above, we know that $\mathbb{P}(\text{Disagreement}) = 1 - p_Y(-1) = 2p(1-p)$. Recall from our tutorials that the expected value of a Geometric(q) random variable is $1/q$. So the expected number of *pairs* of flips is $1/[2p(1-p)]$. By the linearity of expectation, the expected number of *total flips* is $2/[2p(1-p)] = 1/p(1-p)$. This agrees with our intuition from above. The smallest possible value of the expectation is 4, corresponding to $p = 1/2$. As p gets close to zero or one, the expectation number of flips required becomes *arbitrarily large*. So this idea will work best if the coin isn't *too biased*. Otherwise Frank and Ruth will spend all their time flipping coins and never get around to playing tennis!

7. Let X be a continuous RV with support $[0, 1]$ and density function $f_X(x) = 2x$ for $x \in [0, 1]$, zero otherwise.

5 (a) Find the CDF F_X of X.

Solution: For $x \in [0,1]$ we have

$$F_X(x) \equiv \int_{-\infty}^{x} f(t)\,\mathrm{d}t = \int_0^x 2t\,\mathrm{d}t = t^2\big|_0^x = x^2.$$

Therefore,

$$F(x) = \begin{cases} 0, & x < 0 \\ x^2, & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}$$

5 (b) Calculate $\mu_X \equiv \mathbb{E}(X)$.

Solution:

$$\mathbb{E}[X] \equiv \int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x = \int_0^1 2x^2\,\mathrm{d}x = \frac{2}{3}x^3\Big|_0^1 = \frac{2}{3}$$

You are bidding on a government contract that has a value of v to your firm. Suppose that v is measured in "millions of dollars" and $0 \le v \le 1$. Note that v is a known constant from your perspective. If you bid b and your bid is accepted, your firm earns a profit of $\pi = v - b$. Other firms will also bid on this contract, and their values may not equals yours Each firm submits its bid in a sealed envelope. The highest bidder wins the contract and pays their bid; no other bidders make payments.

Suppose that you always bid to maximize your firm's *expected profit*. To choose the optimal bid $b^*(v)$ as a function of your value v, you need to consider the *maximum bid* $B_{\max}$ that your competitors will submit. From your perspective this is a random variable, since you are not party to other firms' cost structures. Based on your research you conclude that $B_{\max}$ has the same CDF as X from above.

5 (c) Derive an expression for your expected profit $\mathbb{E}[\pi(b)]$ as a function of your bid b and your value v.

Solution: Suppose you bid b. There are two cases:

(i) If $b < B_{\max}$, you **Lose** the auction and earn a profit of zero.

(ii) If $b > B_{\max}$, you **Win** the auction and earn a profit of $v - b$.

Since $B_{\max}$ is a continuous random variable $\mathbb{P}(B_{\max} = b) = 0$. This means we can ignore the case of $b = B_{\mathrm{Max}}$. Therefore, by the law of iterated

expectations, your expected profit is

$$\begin{aligned}\mathbb{E}[\pi(b)] &= \mathbb{E}[\pi(b)|\text{Lose}]\mathbb{P}(\text{Lose}) + \mathbb{E}[\pi(b)|\text{Win}]\mathbb{P}(\text{Win}) \\ &= 0 \times \mathbb{P}(\text{Lose}) + (v-b)\mathbb{P}(\text{Win}) \\ &= (v-b)\mathbb{P}(b > B_{\max}).\end{aligned}$$

Again, $\mathbb{P}(B_{\max} = b) = 0$ so $\mathbb{P}(\text{Win}) = \mathbb{P}(B_{\max} \leq b)$. In other words, $\mathbb{E}[\pi(b)] = (v-b)F(b)$ where F is the CDF of $B_{\max}$. Therefore,

$$\mathbb{E}\left[\pi(b)\right] = (v-b)F_X(b) = (v-b)b^2 = vb^2 - b^3$$

by substituting the CDF of X from above: $F_X(x) = x^2$.

5 (d) Using the result of the preceding part, derive a function for your optimal bid $b^*(v)$ as a function of your value v.

Solution: The expected profit function is $\mathbb{E}[\pi(b)] = vb^2 - b^3$. Differentiating with respect to b gives the first-order condition

$$\frac{d}{db}\mathbb{E}\left[\pi(b)\right] = b(2v - 3b) = 0.$$

This is a quadratic with two roots: $b = 0$ and $b = 2v/3$. Intuitively, $b = 0$ cannot be optimal if $v > 0$. If the project has positive value to the firm, there must be *some positive bid* worth paying. In contrast, $b = 2v/3$ looks more plausible. It implies that as your value increase, so should your bid. At the same time, it implies that you should "shade your bid", i.e. bid less than your value.

To confirm that the optimal bidding rule is indeed $b^*(v) = 2v/3$ requires a bit more work. First we check the endpoints: $\mathbb{E}[\pi(0)] = 0$ and $\mathbb{E}[\pi(1)] = v - 1$. Since $0 \leq v \leq 1$, $\mathbb{E}[\pi(1)] \leq 0$. Now suppose $b \neq 0$. Then the first derivative $b(2v - 3b)$ is positive whenever $2v > 3b$ and negative otherwise. Therefore, $\mathbb{E}[\pi(0)]$ is *increasing* for all $b < 2v/3$ and *decreasing* for all $b > 2v/3$. This implies that the maximum is indeed $b^*(v) = 2v/3$.

10 (e) Now suppose that the rules of the auction change. Bids are still submitted in sealed envelopes, the highest bid wins, and only the winner pays. But if you win the auction you pay the bid of the *second highest bidder* rather than your own bid. In other words, if your firm wins you pay $B_{\max}$. Would your optimal bidding rule change? Explain.

Solution: Under this new auction procedure, your expected profit becomes

$$\begin{aligned}\mathbb{E}[\pi(b)] &= 0 \times \mathbb{P}(\text{Lose}) + \mathbb{E}(v - B_{\max}|\text{Win}) \times \mathbb{P}(\text{Win}) \\ &= [v - \mathbb{E}(B_{\max}|B_{\max} < b)]\, b^2.\end{aligned}$$

This is a little tricky since it involves $\mathbb{E}(B_{\max}|B_{\max} < b)$. We need to calculate the expectation of $B_{\max}$ restricted to the interval where $B_{\max} < b$. But how do we do this? We need a probability density function that has the same *shape* as f_X but equals zero outside the interval $[0, b]$. In other words, this new pdf needs to integrate to one between 0 and b. From the CDF of X, we know that the integral of f_X between 0 and b is simply b^2. So if we *divide* f_X by b^2, we'll get something that integrates to one. In other words:

$$\mathbb{E}(B_{\max}|B_{\max} < b) = \int_0^b x \frac{f_X(x)}{b^2}\,\mathrm{d}x = \frac{1}{b^2}\int_0^b x f_X(x)\,\mathrm{d}x.$$

Therefore,

$$\begin{aligned}\mathbb{E}[\pi(b)] &= \left[v - \frac{1}{b^2}\int_0^b x f_X(x)\,\mathrm{d}x\right] b^2 = vb^2 - \int_0^b 2x^2\,\mathrm{d}x \\ &= vb^2 - \frac{2}{3}x^3\Big|_0^b = vb^2 - \frac{2b^3}{3}\end{aligned}$$

and the first-order condition becomes $2b(v - b) = 0$. There are two roots: $b = 0$ and $b = v$. The first of these clearly cannot be the solution: it gives zero profit. When $b < v$ the expected profit function is increasing; when $b > v$ it is decreasing. So $b^* = v$ is a global maximum. This rule says that you should *bid your value* rather than bidding *below your value* as above.

Let's check whether this makes sense. Suppose that $B_{\max} = m$. We don't know m. But we do know that when $b > m$ you win and earn a profit of $v - m$ whereas if $b < m$ you lose and earn a profit of zero. The key point here is that *not matter what you bid* the amount that you pay when you win is always m. There are two cases. First, if $m > v$ then you don't want to win since you'd make a negative profit. Second, if $m < v$ then you want to win, since you'd make a positive profit. But you have no control over the *size* of the profit: it's $(v - m) > 0$ regardless of what you bid. This means that the best possible bidding rule would be one that ensures you *never win* when winning would incur a loss and you *always win* when winning would give you a positive profit. The rule $b^* = v$ achieves exactly this: if $m > v$ it ensures that you lose, and if $m < v$ it ensures that you win. This is another way of reasoning to the solution without all the math from above.