MA2116 21/22 S1 Worked Solutions

Written 20 November 2024, updated 27 November 2024

Warning: While I've done my best to make sure that these answers are correct, I make mistakes! So tread carefully.

Problem 1. *Let X_1, …, X_9 be nine independent identically distributed random variables taking only positive values. Find*

$$\mathbf{E}\Big(\frac{X_1+X_2+X_3}{X_1+\cdots+X_9}\Big) \quad \textit{and} \quad \mathbf{E}\Big(\frac{X_1+\cdots+X_7}{X_1+\cdots+X_9}\Big).$$

Solution. By symmetry all the nine values

$$\mathbf{E}\Big(\frac{X_j}{X_1+\cdots+X_9}\Big)$$

must be equal. Since their sum is 1, they must each have value $1/9$. Thus the desired expectations are $3/9 = 0.\overline{3}$ and $7/9 = 0.\overline{7}$ respectively. Alternatively, since the question suggests that these expectations must have the same values for any choice of X_i satisfying the conditions, we can let each X_i be the constant random variable taking value 1.

Problem 2. *Let Z be a standard normal random variable. For any $a \in \mathbf{R}$, define X_a by*

$$X_a := \begin{cases} Z & \text{if } Z > a, \\ 0 & \text{otherwise.} \end{cases}$$

Find $\mathbf{E}X_0$ and $\mathbf{E}X_1$.

Solution. Since the pdf of Z is $\frac{1}{\sqrt{2\pi}}\exp(-x^2/2)$, we have

$$\mathbf{E}X_a = \int_a^\infty x\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx = -\frac{1}{\sqrt{2\pi}}e^{-a^2/2},$$

which gives $\mathbf{E}X_0 = 1/\sqrt{2\pi} = 0.39894228\ldots$ and $\mathbf{E}X_1 = 1/\sqrt{2e\pi} = 0.24197\ldots$.

Problem 3. *Three numbers A, B, and C are selected independently at random from the unit interval $[0,1]$. What is the probability that both roots of the equation $Ax^2+Bx+C=0$ are real?*

Solution. This amounts to finding $\Pr(B^2 \geq 4AC)$. We approach this by finding the distributions of B^2 and AC first. Write

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x > 1 \end{cases}$$

for the CDF of A, B, and C. Then for $0 \leq t \leq 1$ we have

$$\Pr(B^2 \leq t) = \Pr(B \leq \sqrt{t}) = \begin{cases} 0 & \text{if } t < 0 \\ \sqrt{t} & \text{if } 0 \leq t \leq 1 \\ 1 & \text{if } t > 1, \end{cases}$$

which yields the probability density function (by differentiation)

$$f_{B^2}(t) = \frac{1}{2\sqrt{t}} 1_{[0,1]}(t).$$

For the CDF of AC we use the law of total probability to compute

$$\begin{aligned}
\Pr(AC \le t) &= \int_{-\infty}^{\infty} \Pr(AC \le t \mid C = x) f_C(x)\, dx \\
&= \int_{-\infty}^{\infty} \Pr(A \le t/x) 1_{[0,1]}(x)\, dx \\
&= \int_0^1 F(t/x)\, dx \\
&= \int_0^t 1\, dx + \int_t^1 \frac{t}{x}\, dx \\
&= t - t\log t.
\end{aligned}$$

Consequently by the law of total probability again we have

$$\begin{aligned}
\Pr(4AC \le B^2) &= \int_{-\infty}^{\infty} \Pr(4AC \le t) f_{B^2}(t)\, dt \\
&= \int_0^1 \left(\frac{t}{4} - \frac{t}{4}\log\frac{t}{4}\right) \frac{1}{2\sqrt{t}}\, dt \\
&= \frac{1}{36}(5 + \log 64) \\
&= 0.254413419\ldots.
\end{aligned}$$

Problem 4. *Let X be a Poisson random variable with mean 10. What is the upper bound of the probability of $X \ge 12$ given by (1) Markov's inequality? (2) the one-sided Chebyshev inequality? (3) the Chernoff bound? (4) What is the approximate value of the probability of $X \ge 12$ given by the central limit theorem?*

Solution. (1) Recall that *Markov's inequality* states that $\Pr(X \ge a) \le \mathbf{E}(X)/a$ whenever X is a nonnegative random variable and $a > 0$. We have $\mathbf{E}X = 10$, so Markov's inequality gives $\Pr(X \ge 12) \le 10/12 = 0.8\overline{3}$.

(2) Recall that the *one-sided Chebyshev inequality* states that $\Pr(X - \mathbf{E}X \ge a) \le \frac{\sigma^2}{\sigma^2 + a^2}$, where $\sigma^2 < +\infty$ is the variance of X. We have $\mathbf{E}X = \sigma^2 = 10$, so we have $\Pr(X - 10 \ge 2) \le \frac{10}{10+4} = 0.\overline{714285}$.

(3) Recall that the *Chernoff bound* states that $\Pr(X \ge a) \le e^{-ta}\mathbf{E}e^{tX}$ for all $t > 0$, where $\mathbf{E}e^{tX}$ is the moment generating function of X. Since $\mathbf{E}e^{tX} = e^{10(e^t - 1)}$, we have

$$\Pr(X \ge 12) \le e^{-12t} e^{10(e^t-1)} = e^{2(-5+5e^t-6t)}$$

for all $t > 0$. To minimize this quantity we need only to minimize $5e^t - 6t$ owing to the fact that $\exp$ is increasing on $\mathbf{R}$. Differentiation reveals that the global minimum is achieved at $\log(6/5)$. Substituting this value gives the bound

$$\Pr(X \ge 12) \le \exp\big(2 - 12\log(6/5)\big) = 0.828731814\ldots.$$

(4) See Example 3b in Chapter 8 of Ross, *A First Course in Probability* (10th edition). The idea is to treat the Poisson random variable with mean 10 as the sum of 10 independent Poisson random variables each with mean 1. Then the central limit theorem gives

$$
\begin{aligned}
\Pr(X \geq 12) &= \Pr(X \geq 11.5) \\
&= \Pr\left(\frac{X-10}{\sqrt{10}} \geq \frac{11.5-10}{\sqrt{10}}\right) \\
&\approx 1 - \Phi\left(\frac{1.5}{\sqrt{10}}\right) \\
&\approx 0.3176281479986.
\end{aligned}
$$

Note that the actual value of $\Pr(X \geq 12)$ is around 0.303224.

Problem 5. *A new test for COVID is designed to work by measuring the concentration of antigens of the virus in a standard nasal sample. The concentration of antigens is assumed to follow a normal distribution:*

- *with mean* 7.0 *and variance* 2.0 *among COVID-infected people, but*
- *with mean* 6.0 *and variance* 1.5 *among COVID-uninfected (healthy) people.*

The test is calibrated to indicate:

- *positive (for COVID) if the measured concentration is* ≥ 6.5*, and*
- *negative (for COVID) if the measured concentration is* < 6.5*.*

(1) What is the sensitivity of the test? (2) What is the specificity of the test?

Solution. (See L1 Lesson 11 Problem 2.) (1) The sensitivity (true positive rate) is given by $\Pr(N(\mu = 7, \sigma^2 = 2) \geq 6.5) = 0.638163195\ldots$ and (2) the specificity (true negative rate) is given by $\Pr(N(\mu = 6, \sigma^2 = 1.5) < 6.5) \approx 0.65845430$.

Problem 6. *A breath-analyzer test reports a sensitivity of* 95% *and a specificity of* 97% *for detecting drink-driving (blood alcohol content exceeding the legal limit). The prevalence of drink-driving in the population is* 1%*. A driver is stopped at random and given the breath-analyzer test. The test gives a positive result. What is the probability that the driver is drink-driving?*

Solution. Write DD for drunk-driving, and NDD for not drunk-driving. We are given $\Pr(DD) = 0.01$, $\Pr(+ \mid DD) = 0.95$, and $\Pr(- \mid NDD) = 0.97$, from which we may compute (via Bayes's theorem and the law of total probability)

$$
\begin{aligned}
\Pr(DD \mid +) &= \frac{\Pr(+ \mid DD)\Pr(DD)}{\Pr(+ \mid DD)\Pr(DD) + \Pr(+ \mid NDD)\Pr(NDD)} \\
&= \frac{0.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.03 \cdot 0.99} \\
&= \frac{95}{392} \approx 0.2423469.
\end{aligned}
$$

Problem 7. *A box contains*

- 3 *coins of type* A*, each biased with a probability of* 80% *showing heads,*
- 2 *coins of type* B*, each biased with a probability of* 55% *showing heads, and*
- 1 *coin of type* C*, biased with a probability of* 30% *showing heads.*

One of these six coins in the box is chosen at random and is flipped 10 *times. Given that exactly* 2 *heads appear among the first* 4 *flips, (1) what is the expected number of heads obtained in the* 10 *flips? (2) what is the probability of getting at least* 7 *heads in the* 10 *flips?*

Solution. (1) Write X for the number of heads that appear in the 10 flips, and Y for the number of heads appearing among the first 4 flips. Our task is to compute $\mathbf{E}(X \mid Y=2)$, which is given by the law of total probability

$$\sum_i \mathbf{E}\big((X \mid T_i) \mid Y=2\big)\Pr(T_i \mid Y=2),$$

the sum ranging over the three coin types $(T_1, T_2, T_3) = (A, B, C)$.

The first step is to find $\Pr(Y=2)$. We have $\Pr(A) = 3/6 = 1/2$, $\Pr(B) = 2/6 = 1/3$, and $\Pr(C) = 1/6$. Since $\Pr(Y=2 \mid A) = \binom{4}{2}(0.8)^2(0.2)^2 = 0.1536$,
$\Pr(Y=2 \mid B) = \binom{4}{2}(0.55)^2(0.45)^2 = 29403/80000$, and
$\Pr(Y=2 \mid C) = \binom{4}{2}(0.3)^2(0.7)^2 = 0.2646$, we may use Bayes's theorem to get

$$\begin{aligned}\Pr(Y=2) &= \Pr(Y=2 \mid A)\Pr(A) + \Pr(Y=2 \mid B)\Pr(B) + \Pr(Y=2 \mid C)\Pr(C) \\ &= 0.2434.\end{aligned}$$

Now we can use this to compute the values $\Pr(T_i \mid Y=2)$. We have

$$\Pr(A \mid Y=2) = \frac{\Pr(Y=2 \mid A)\Pr(A)}{\Pr(Y=2)} = \frac{2048}{6491},$$

and similarly $\Pr(B \mid Y=2) = 3267/6491$ and $\Pr(C \mid Y=2) = 1176/6491$. Finally the individual conditional expectations are straightforward: For example, we have
$\mathbf{E}\big((X \mid A) \mid Y=2\big) = 2 + 6 \cdot 0.8 = 6.8$; similarly
$\mathbf{E}\big((X \mid B) \mid Y=2\big) = 2 + 6 \cdot 0.55 = 5.3$ and $\mathbf{E}\big((X \mid C) \mid Y=2\big) = 3.8$. Putting this all together gives

$$\mathbf{E}(X \mid Y=2) = \frac{357103}{64910} = 5.501509782776\ldots.$$

(2) We want $\Pr(X \geq 7 \mid Y=2)$, and this is given much like before by the law of total probability

$$\sum_i \Pr\big((X \geq 7 \mid T_i) \mid Y=2\big)\Pr(T_i \mid Y=2).$$

We have the probabilities $\Pr(T_i \mid Y=2)$ from earlier. For the former probabilities, we have for example $\Pr\big((X \geq 7 \mid A) \mid Y=2\big) = \Pr(Z \geq 5)$ with $Z \sim \text{Binomial}(6, 0.8)$; the three values for $\sum_i \Pr\big((X \geq 7 \mid T_i) \mid Y=2\big)$ are then

$$\frac{2048}{3125}, \quad \frac{2093663}{12800000}, \quad \text{and} \quad \frac{2187}{200000}$$

for $(T_1, T_2, T_3) = (A, B, C)$ respectively, which gives our final probability as

$$\Pr(X \geq 7 \mid Y=2) = \frac{24\,184\,468\,573}{83\,084\,800\,000} \approx 0.291081745.$$

(In general for such questions one treats $\Pr(\cdot \mid)$ and $\mathbf{E}(\cdot \mid A)$ as replacements for $\Pr(\cdot)$ and $\mathbf{E}(\cdot)$, thinking of this as us passing to a new world A where we have integrated the knowledge of the event A and how it affects all our probabilities. Thinking this way makes it easy to remember the appropriate analogues of the usual unconditioned laws of total probability and expectation.)

Problem 8. *An urn contains:*

- 3 *red balls*

- 4 *blue balls*
- 5 *white balls*

We selected 6 *balls at random from this urn. The random variables* X *and* Y *denote respectively the number of red and blue balls among those selected. Find (1)* $\mathbf{E}(XY)$ *and (2)* $\mathrm{Cov}(X,Y)$.

Solution. (1) For $1 \le j \le 6$, write X_j for the indicator random variable of the event where the jth ball selected is red; similarly Y_j for when the jth ball selected is blue. Then $X = \sum X_j$, $Y = \sum Y_j$, and

$$\mathbf{E}(XY) = \sum_{1\le i,j\le 6} \mathbf{E}(X_iY_j) = \sum_{1\le i\ne j\le 6} \Pr(X_iY_j).$$

Now $\Pr(X_iY_j) = \frac{3\cdot 4}{12\cdot 11}$, so $\mathbf{E}(XY) = 30\Pr(X_iY_j) = 30/11 = 2.\overline{72}$.

(2) Recall that $\mathrm{Cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$. We have $\mathbf{E}X = \sum \mathbf{E}X_i = \sum \Pr(X_i)$, and $\Pr(X_i) = 3/12$, so $\mathbf{E}X = 6\Pr(X_i) = 3/2$. Similarly $\mathbf{E}Y = 6\cdot 4/12 = 2$, so $\mathrm{Cov}(X,Y) = 30/11 - 3 = -3/11 = -0.\overline{27}$.

Problem 9. *Let* X *and* Y *be independent random variables uniformly distributed on the unit interval* $[0,1]$. *Find (1)* $\Pr(-0.5 < 3X - 2Y < 0.5)$ *and (2)* $\Pr(0 < 3X - 2Y < 2.5)$.

Solution. (See L1 Lesson 16 Problem 1.) We compute the probability density function f of $3X - 2Y$. We have $f_X = f_Y = 1_{[0,1]}$, $f_{3X} = \frac{1}{3}1_{[0,3]}$, and $f_{-2Y} = \frac{1}{2}1_{[-2,0]}$, so f is given by the convolution

$$\begin{aligned} f(a) = (f_{3X} * f_{-2Y})(a) &= \int_{-\infty}^{\infty} f_{3X}(a-y) f_{-2Y}(y)\,dy \\ &= \frac{1}{6}\int_{-\infty}^{\infty} 1_{[0,3]}(a-y)1_{[-2,0]}(y)\,dy \\ &= \frac{1}{6}\int_{a-3}^{a} 1_{[-2,0]}(y)\,dy \\ &= \frac{1}{6}\cdot\begin{cases} a+2 & \text{if } -2\le a\le 0; \\ 2 & \text{if } 0\le a\le 1; \\ 3-a & \text{if } 1\le a\le 3; \\ 0 & \text{otherwise.}\end{cases} \end{aligned}$$

(1) We have

$$\begin{aligned} \Pr(-0.5 < 3X - 2Y < 0.5) &= \int_{-0.5}^{0.5} f(a)\,da \\ &= \frac{1}{6}\int_{-0.5}^{0} (a+2)\,da + \frac{1}{6}\int_0^{0.5} 2\,da \\ &= \frac{5}{16} = 0.3125. \end{aligned}$$

(2) We have

$$\Pr(0 < 3X - 2Y < 2.5) = \frac{1}{6}\int_0^1 2\,da + \frac{1}{6}\int_1^{2.5}(3-a)\,da = \frac{31}{48} = 0.6458\overline{3}.$$

Problem 10. *Suppose the random variables* X *and* Y *have a joint density function given by*

$$f(x,y) = \begin{cases} 8/(x^3y^5) & \text{if } x,y \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find (1) $\Pr(0.5 < XY < 1.5 \text{ and } 0.5 < X/Y < 1.5)$ *and (2)* $\Pr(2 < XY < 3 \text{ and } 1 < X/Y < 2)$.

Solution. (See L1 Lesson 18 Problem 1.) We perform a change of variables, writing $U = XY, V = X/Y$, and $\Phi(x,y) = (xy, x/y)$. Note that $x = \sqrt{uv}$ and $y = \sqrt{u/v}$ (this is valid on the region $x, y \geq 1$). The Jacobian is given by

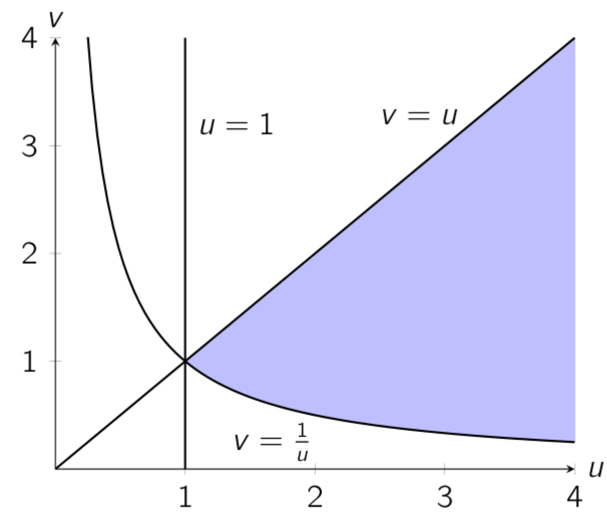
$$|J_\Phi| = \det \begin{pmatrix} y & x \\ 1/y & -x/y^2 \end{pmatrix} = -\frac{2x}{y} = -\frac{2\sqrt{uv}}{\sqrt{u/v}}.$$

The joint density function in the new coordinates is then given by

$$\begin{aligned} g(u,v) = f(x,y) &= f(\Phi^{-1}(u,v))|J_\Phi|^{-1} \\ &= \frac{8}{(\sqrt{uv})^3(\sqrt{u/v})^5}\frac{\sqrt{u/v}}{2\sqrt{uv}} = \frac{4}{u^4} \end{aligned}$$

over the region

$$\left\{(u,v) \;\middle|\; \sqrt{uv} \geq 1 \text{ and } \sqrt{u/v} \geq 1\right\} = \left\{(u,v) \;\middle|\; u \geq 1 \text{ and } 1/u \leq v \leq u\right\}.$$



(1) We have

$$\Pr(0.5 < XY < 1.5 \text{ and } 0.5 < X/Y < 1.5) = \int_1^{1.5}\int_{1/u}^{u} \frac{4}{u^4}\,dv\,du = \frac{25}{81} \approx 0.3086419753.$$

(2) We have

$$\Pr(2 < XY < 3 \text{ and } 1 < X/Y < 2) = \int_2^3\int_1^2 \frac{4}{u^4}\,dv\,du = \frac{19}{162} \approx 0.11728395.$$

Problem 11. *Suppose* X *and* Y *are jointly continuous random variables whose joint density function is given by*

$$f(x,y) = \begin{cases} 2 & \text{if } 0 < x < y < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Find (1) $\Pr(1/4 < X < 1/2 \mid Y = 1/2)$, *(2)* $\Pr(1/4 < Y < 1/2 \mid X = 1/2)$, *(3)* $\mathbf{E}(XY)$, *and (4)* $\operatorname{Cov}(X,Y)$.

Solution. (See L1 Lesson 17 Problem 1.) (1) To find this probability, we need the conditional density function $f_{X|Y}(x \mid y) = f(x,y)/f_Y(y)$. So we first compute the marginal density function

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^y 2\,dx = 2y$$

for $0 < y < 1$, and $f_Y(y) = 0$ otherwise. Then $f_{X|Y}(x \mid y) = f(x,y)/f_Y(y) = 1/y$ for $0 < x < y < 1$, so

$$\Pr(1/4 < X < 1/2 \mid Y = 1/2) = \int_{1/4}^{1/2} \frac{1}{(1/2)}\,dx = \frac{1}{2}.$$

(2) No computation is needed: This probability is 0, since it requires $Y < 1/2$ and $X = 1/2$, whereas f is nonzero only when $y > x$.

(3) We have

$$\mathbf{E}(XY) = \int_0^1 \int_x^1 2xy\,dy\,dx = \frac{1}{4}.$$

(4) We have $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_x^1 2\,dy = 2 - 2x$ for $0 < x < 1$ and 0 otherwise, so $\mathbf{E}X = \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_0^1 (2x - 2x^2)\,dx = 1/3$. Similarly $\mathbf{E}Y = \int_{-\infty}^{\infty} y f_Y(y)\,dy = \int_0^1 2y^2 = 2/3$. It follows that $\operatorname{Cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y) = 1/36 = 0.02\overline{7}$.

Problem 12. *We are given a coin. It can be tossed any number of times to generate independent outcomes (heads or tails). Its probability X of showing heads is initially assumed to be uniformly distributed on $(0,1)$. We toss the coin 4 times and see exactly 2 heads among them. Before any further toss of the coin is made:*

1. *What would be our best prediction on the number of heads occurring when it is tossed another 4 times?*
2. *What is the mean squared error of our best prediction?*

We now toss the coin another 4 times, and see exactly 1 head among them:

3. *What would now be our best prediction on the number of heads occurring when the coin is tossed another 4 times?*
4. *What is the mean squared error of our new prediction?*

A friend now joins us and is only told that after tossing the coin a total of 8 times, we saw exactly 3 heads among them:

5. *What would be our friend's best prediction on the number of heads occurring when the coin is tossed another 4 times?*
6. *What is the mean squared error of that prediction?*

Solution. (See L1 Lesson 23 Problem 1.) (1) Write N for the number of heads occurring when the coin is tossed another 4 times. Our best prediction is $\mathbf{E}N$. Intuitively the answer is 2, as an average guess with no information is that the coin is fair, and the initial experiment we conduct gives us no information that suggests otherwise. We can make this rigorous as follows: We have $f_X(p) = 30p^2(1-p)^2$ (see Example 5e in chapter 6 of Ross 10th edition). Since $N|(X = p) \sim \text{Binomial}(4, p)$, we have for $0 \leq k \leq 4$

$$\begin{aligned}\Pr(N = k) &= \int_{-\infty}^{\infty} \Pr(N = k \mid X = p) f_X(p)\, dp \\ &= \int_0^1 \binom{4}{k} p^k (1-p)^{4-k} \cdot 30p^2(1-p)^2\, dp \\ &= 30\binom{4}{k}\, \mathrm{B}(k+3, 7-k).\end{aligned}$$

(Here we have used the beta function, which satisfies the identity $\mathrm{B}(m, n) = \int_0^1 t^{m-1}(1-t)^{n-1}\, dt = (m-1)!(n-1)!/(m+n-1)!$.) This takes values $(5/42, 5/21, 2/7, 5/21, 5/42)$ for $k = (0, 1, 2, 3, 4)$, and this corresponds intuitively to the fact that the observation of two heads in four flips gives us some confidence in believing the coin is more likely to be fair than not. It is straightfoward to compute then that $\mathbf{E}N = 2$.

(2) This is the variance $\mathbf{E}((N-\mu)^2) = \sum_{k=0}^{4} (k-2)^2 \cdot p_k = 10/7$.

(3) (See Example 5e in chapter 6 of Ross 10ed.) Write N' for the new number of heads occurring in 4 tosses, and X' for the new distribution of X conditioned on the event just observed, that 4 trials (tosses) yield 1 success (heads). We then have the conditional density

$$\begin{aligned}f_{X'}(x) &= \frac{\Pr\left(\substack{\text{4 tosses yields 1 head} \\ \text{given that } X = x}\right) f_X(x)}{\Pr(\text{4 tosses yields 1 head})} \\ &= \frac{\binom{4}{1} x(1-x)^3}{\Pr(\text{4 tosses yields 1 head})} 1_{[0,1]}(x) \\ &= cx(1-x)^3;\end{aligned}$$

in general if the observation yielded n successes in $n+m$ trials we would get a new conditional density $cx^n(1-x)^m$. In fact $c = 20 = 1/\mathrm{B}(1+1, 3+1)$.

Finally note that $N'|(X' = x) \sim \text{Binomial}(4, x)$, so its expectation is simply $4x$. Now we get

$$\begin{aligned}\mathbf{E}N' &= \int_{-\infty}^{\infty} \mathbf{E}(N' \mid X' = x) f_{X'}(x)\, dx \\ &= \int_0^1 4x \cdot 20x(1-x)^3\, dx \\ &= 80B(3, 4) = \frac{4}{3} = 1.\overline{3}.\end{aligned}$$

(This passes a sanity check: Without any information I would have guessed 2 heads in 4 throws of a coin, but the information that 4 tosses yielded only 1 head causes me to lower my guess to somewhere between 1 and 2.)

(4) We compute $\mathbf{E}N'^2$ first. We have $\text{Var}(N' \mid X' = x) = 4x(1-x)$, so $\mathbf{E}(N'^2 \mid X' = x) = 4x(1-x) + (4x)^2 = 4x + 12x^2$. Then

$$
\begin{aligned}
\mathbf{E}(N'^2) &= \int_{-\infty}^{\infty} \mathbf{E}(N'^2 \mid X' = x) f_{X'}(x)\,dx \\
&= \int_0^1 (4x + 12x^2) \cdot 20x(1-x)^3\,dx \\
&= \int_0^1 80x^2 - 480x^4 + 640x^5 - 240x^6\,dx \\
&= \frac{64}{21}.
\end{aligned}
$$

Thus $\mathrm{Var}(N'^2) = (64/21) - (4/3)^2 = 80/63 \approx 1.26984$. (Warning: $\mathrm{Var}(N) \neq \int_{-\infty}^{\infty} \mathrm{Var}(N \mid X = x) f_X(x)\,dx$ in general! The correct formula is more complicated (lookup Law of Total Variance)).

(5) The ideas are the same as in (3) and (4), so we will be brief. Write N'' for the new number of heads occurring in 4 tosses, and X'' for the new distribution of X conditioned on the event just observed, that 8 trials (tosses) yield 3 success (heads). We then have the conditional density $f_{X''}(x) = 504x^3(1-x)^5$ for $0 \leq x \leq 1$, and our friend's best prediction is then

$$
\begin{aligned}
\mathbf{E}N'' &= \int_{-\infty}^{\infty} \mathbf{E}(N'' \mid X'' = x) f_{X''}(x)\,dx \\
&= \int_0^1 4x \cdot 504x^3(1-x)^5\,dx \\
&= \frac{8}{5} = 1.6.
\end{aligned}
$$

(6) As in (4) we have $\mathbf{E}(N''^2 \mid X'' = x) = 4x + 12x^2$, and so

$$
\begin{aligned}
\mathbf{E}(N''^2) &= \int_{-\infty}^{\infty} \mathbf{E}(N''^2 \mid X'' = x) f_{X''}(x)\,dx \\
&= \int_0^1 (4x + 12x^2) \cdot 504x^3(1-x)^5\,dx \\
&= \frac{208}{55} = 3.7\overline{81}.
\end{aligned}
$$

It follows that the mean squared error or variance is given by $(208/55) - (8/5)^2 = 336/275 = 1.22\overline{18}$.