Probability Review for Final Exam
18.05 Spring 2014
Jeremy Orloff and Jonathan Bloom

Problem 1. **(a)** Four ways to fill each slot: $\boxed{4^5}$.

(b) Four ways to fill the first slot and 3 ways to fill each subsequsent slot: $\boxed{4 \cdot 3^4}$.

(c) Build the sequences as follows:
Step 1: Choose which of the 5 slots gets the A: 5 ways to place the one A.
Step 2: 3^4 ways to fill the remain 4 slots. By the rule of product there are $\boxed{5 \cdot 3^4}$ such sequences.

Problem 2. **(a)** $\binom{52}{5}$.

(b) Number of ways to get a full-house: $\binom{4}{2}\binom{13}{1}\binom{4}{3}\binom{12}{1}$

(c) $\dfrac{\binom{4}{2}\binom{13}{1}\binom{4}{3}\binom{12}{1}}{\binom{52}{5}}$

Problem 3. There are several ways to think about this. Here is one.

The 11 letters are p, r, o, b,b, a, i,i, l, t, y. We use the following steps to create a sequence of these letters.
Step 1: Choose a position for the letter p: 11 ways to do this.
Step 2: Choose a position for the letter r: 10 ways to do this.
Step 3: Choose a position for the letter o: 9 ways to do this.
Step 4: Choose two positions for the two b's: 8 choose 2 ways to do this.
Step 5: Choose a position for the letter a: 6 ways to do this.
Step 6: Choose two positions for the two i's: 5 choose 2 ways to do this.
Step 7: Choose a position for the letter l: 3 ways to do this.
Step 8: Choose a position for the letter t: 2 ways to do this.
Step 9: Choose a position for the letter y: 1 ways to do this.

Multiply these all together we get:

$$11 \cdot 10 \cdot 9 \cdot \binom{8}{2} \cdot 6 \cdot \binom{5}{2} \cdot 3 \cdot 2 \cdot 1 \;=\; \frac{11!}{2! \cdot 2!}$$

Problem 4. We are given $P(E \cup F) = 3/4$.
$E^c \cap F^c = (E \cup F)^c \;\Rightarrow\; P(E^c \cap F^c) = 1 - P(E \cup F) = \boxed{1/4.}$

Problem 5. D is the disjoint union of $D \cap C$ and $D \cap C^c$.
So, $P(D \cap C) + P(D \cap C^c) = P(D)$
$\Rightarrow\ P(D \cap C^c) = P(D) - P(D \cap C) = .4 - .2 = \boxed{.2.}$

Problem 6. **(a)** $A = \{\text{HTT, THT, TTH}\}$.
$B = \{\text{HTT, THT, TTH, TTT}\}$.
$C = \{\text{HHH, HHT, HTH, HTT, TTT}\}$
(There is some ambiguity here, we'll also accept $C = \{\text{HHH, HHT, HTH, HTT}\}$)
$D = \{\text{THH, THT, TTH, TTT}\}$.
(b) $A^c = \{\text{HHH, HHT, HTH, THH, TTT}\}$
$A \cup (C \cap D) = \{\text{HTT, THT, TTH, TTT}\}$. (Also accept $\{\text{HTT, THT, TTH}\}$.)
$A \cap D^c = \{\text{HTT}\}$.

Problem 7. **(a)** Slots 1, 3, 5, 7 are filled by T_1, T_3, T_5, T_7 in any order: 4! ways.
Slots 2, 4, 6, 8 are filled by T_2, T_4, T_6, T_8 in any order: 4! ways.
answer: $4! \cdot 4! = 576$.
(b) There are 8! ways to fill the 8 slots in any way.
Since each outcome is equally likely the probabilitiy is
$\dfrac{4! \cdot 4!}{8!} = \dfrac{576}{40320} = 0.143 = 1.43\%$.

Problem 8. Let H_i be the event that the i^{th} hand has one king. We have the conditional probabilities

$$P(H_1) = \frac{\binom{4}{1}\binom{48}{12}}{\binom{52}{13}}; \quad P(H_2|H_1) = \frac{\binom{3}{1}\binom{36}{12}}{\binom{39}{13}}; \quad P(H_3|H_1 \cap H_2) = \frac{\binom{2}{1}\binom{24}{12}}{\binom{26}{13}}$$

$P(H_4|H_1 \cap H_2 \cap H_3) = 1$

$$\begin{aligned} P(H_1 \cap H_2 \cap H_3 \cap H_4) &= P(H_4|H_1 \cap H_2 \cap H_3)\, P(H_3|H_1 \cap H_2)\, P(H_2|H_1)\, P(H_1) \\ &= \frac{\binom{2}{1}\binom{24}{12}\binom{3}{1}\binom{36}{12}\binom{4}{1}\binom{48}{12}}{\binom{26}{13}\binom{39}{13}\binom{52}{13}}. \end{aligned}$$

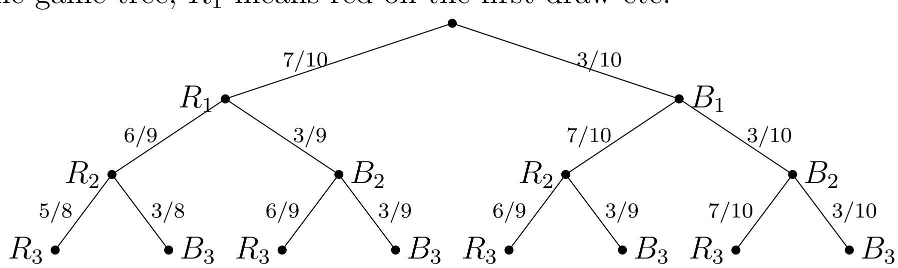
Problem 9. Sample space $= \Omega = \{(1,1),\ (1,2),\ (1,3),\ \ldots,\ (6,6)\,\} = \{(i,j)\,|\,i,j = 1,\ 2,\ 3,\ 4,\ 5,\ 6\,\}$.
(Each outcome is equally likely, with probability 1/36.)

$A = \{(1,3),\ (2,2),\ (3,1)\}$,
$B = \{(3,1),\ (3,2),\ (3,3),\ (3,4),\ (3,5),\ (3,6),\ (1,3),\ (2,3),\ (4,3),\ (5,3),\ (6,3)\}$
$P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{2/36}{11/36} = \dfrac{2}{11}$..
(b) $P(A) = 3/36 \neq P(A|B)$, so they are not independent.

Problem 10. We compute all the pieces needed to apply Bayes' rule.
We're given $P(T|B) = .7 \Rightarrow P(T^c|B) = .3, \quad P(T|B^c) = .1 \Rightarrow P(T^c|B^c) = .9.$
$P(B) = 1.3 \times 10^{-5} \Rightarrow P(B^c) = 1 - P(B) = 1 - 1.3 \times 10^{-5}.$
We use the law of total probability to compute $P(T)$:
$P(T) = P(T|B)\,P(B) + P(T|B^c)\,P(B^c) = .1000078.$
Now we can use Bayes' rule to answer the question:
$P(B|T) = \dfrac{P(T|B)\,P(B)}{P(T)} = 9.10 \times 10^{-5}, \quad P(B|T^c) = \dfrac{P(T^c|B)\,P(B)}{P(T^c)} = 4.33 \times 10^{-6},$

Problem 11. For a given problem let C be the event the student gets the problem correct and K the event the student knows the answer.
The question asks for $P(K|C)$.
We'll compute this using Bayes' rule: $P(K|C) = \dfrac{P(C|K)\,P(K)}{P(C)}$.
We're given: $P(C|K) = 1, \quad P(K) = 0.6.$
Law of total prob.:
$P(C) = P(C|K)\,P(K) + P(C|K^c)\,P(K^c) = 1 \cdot 0.6 + 0.25 \cdot 0.4 = 0.7.$
Therefore $P(K|C) = \dfrac{0.6}{0.7} = .857 = 85.7\%.$

Problem 12.
Here is the game tree, R_1 means red on the first draw etc.



Summing the probability to all the B_3 nodes we get
$P(B_3) = \dfrac{7}{10} \cdot \dfrac{6}{9} \cdot \dfrac{3}{8} + \dfrac{7}{10} \cdot \dfrac{3}{9} \cdot \dfrac{3}{9} + \dfrac{3}{10} \cdot \dfrac{7}{10} \cdot \dfrac{3}{9} + \dfrac{3}{10} \cdot \dfrac{3}{10} \cdot \dfrac{3}{10} = .350.$

Problem 13. We have $P(A \cup B) = 1 - 0.42 = 0.58$ and we know

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Thus,

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.4 + 0.3 - 0.58 = 0.12 = (0.4)(0.3) = P(A)P(B)$$

So A and B are independent.

Problem 14. We have

$$P(A \cap B \cap C) = 0.06 \qquad P(A \cap B) = 0.12$$
$$P(A \cap C) = 0.15 \qquad P(B \cap C) = 0.2$$

Since $P(A \cap B) = P(A \cap B \cap C) + P(A \cap B \cap C^c)$, we find $P(A \cap B \cap C^c) = 0.06$. Similarly

$$\begin{aligned} P(A \cap B \cap C^c) &= 0.06 \\ P(A \cap B^c \cap C) &= 0.09 \\ P(A^c \cap B \cap C) &= 0.14 \end{aligned}$$

Problem 15. **(a)** E = even numbered = {Feb, Apr, Jun, Aug, Oct, Dec}.
F = first half = {Jan, Feb, Mar, Apr, May, Jun}.
S = summer = {Jun, Jul, Aug}.
(a) $E \cap F$ = {Feb, Apr, Jun} $\Rightarrow$ $P(E|F) = 3/6 = P(E)$. So, they are independent.
(b) $E \cap S$ = {Jun, Aug} $\Rightarrow$ $P(E|S) = 2/3 \neq P(E)$. So, they are not independent.

Problem 16. To show A and B are not independent we need to show $P(A \cap B) \neq P(A) \cdot P(B)$.
(a) No, they cannot be independent: $A \cap B = \emptyset \Rightarrow P(A \cap B) = 0 \neq P(A) \cdot P(B)$.
(b) No, they cannot be independent: same reason as in part (a).
(c) No, they cannot be independent: $A \subset B \Rightarrow A \cap B = A$
$\Rightarrow P(A \cap B) = P(A) > P(A) \cdot P(B)$. The last inequality follows because $P(B) < 1$.
(d) No, they cannot be independent: (This one is a little tricky.)
To ease notation, write $P(A) = a$, $P(B) = b$.
A and B are independent implies $P(A \cap B) = ab$.
We know $P(A \cup B) = P(A) + P(B) - P(A \cap B) = a + b - ab$.
Here's the trickiness:

$$a + b - ab = a + b(1 - a) < a + (1 - a) = 1, \quad \text{so} \quad P(A \cup B) < 1.$$

The final bit of the proof is to show that if we assume $A \cup B$ and A are independent we are lead to a contradiction. This is easy, since if they are independent then

$$P(A \cup B|A) = P(A \cup B),$$

but it is clear that $P(A \cup B|A) = 1$.
Since we can't have both $P(A \cup B) < 1$ and $P(A \cup B) = 1$ the assumption of independence must be false.

Problem 17. We compute

$$E[X] = -2 \cdot \frac{1}{15} + -1 \cdot \frac{2}{15} + 0 \cdot \frac{3]15 + 1 \cdot \frac{[}{4}}{15} + 2 \cdot \frac{5}{15} = \frac{2}{3}.$$

Thus

$$\text{Var}(X) = E((X - \frac{2}{3})^2) = \frac{14}{9}.$$

Problem 18. We first compute

$$E[X] = \int_0^1 x \cdot 2x dx = \frac{2}{3}$$

$$E[X^2] = \int_0^1 x^2 \cdot 2x dx = \frac{1}{2}$$

$$E[X^4] = \int_0^1 x^4 \cdot 2x dx = \frac{1}{3}.$$

Thus,

$$\text{Var}(X) = E[X^2] - (E[X])^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

and

$$\text{Var}(X^2) = E[X^4] = \left(E[X^2]\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

Problem 19. **(a)** We have

X values:	-1	0	1
prob:	1/5	2/5	2/5
X^2	1	0	1

So, $\boxed{E(X) = -1/5 + 2/5 = 1/5.}$

(b)

y values:	0	1
prob:	2/5	3/5

$\Rightarrow \boxed{E(Y) = 3/5.}$

(c) The change of variables formula just says to use the bottom row of the table in part (a): $E(X^2) = 1 \cdot (1/5) + 0 \cdot (2/5) + 1 \cdot (2/5) = 3/5$ (same as part (b)).

(d) $\boxed{\text{Var}(X) = E(X^2) - E(X)^2 = 3/5 - 1/25 = 14/25.}$

Problem 20. Use $\text{Var}(X) = E(X^2) - E(X)^2 \Rightarrow 3 = E(X^2) - 4 \Rightarrow \boxed{E(X^2) = 7.}$

Problem 21. **answer:**

Make a table

X:	0	1
prob:	(1-p)	p
X^2	0	1.

From the table, $\boxed{E(X) = 0 \cdot (1-p) + 1 \cdot p = p.}$

Since X and X^2 have the same table $E(X^2) = E(X) = p$.

Therefore, $\boxed{\text{Var}(X) = p - p^2 = p(1-p).}$

Problem 22. Let X be the number of people who get their own hat.

Following the hint: let X_j represent whether person j gets their own hat. That is, $X_j = 1$ if person j gets their hat and 0 if not.

We have, $X = \sum_{j=1}^{100} X_j$, so $E(X) = \sum_{j=1}^{100} E(X_j)$.

Since person j is equally likely to get any hat, we have $P(X_j = 1) = 1/100$. Thus, $X_j \sim \text{Bernoulli}(1/100) \Rightarrow E(X_j) = 1/100 \Rightarrow \boxed{E(X) = 1.}$

Problem 23. For $y = 0, 2, 4, \ldots, 2n$,

$$P(Y = y) = P(X = \frac{y}{2}) = \binom{n}{y/2} \frac{1}{2}^n.$$

Problem 24. We have $f_X(x) = 1$ for $0 \le x \le 1$. The cdf of X is

$$F_X(x) = \int_0^x f_X(t)dt = \int_0^x 1dt = x.$$

Now for $5 \le y \le 7$, we have

$$F_Y(y) = P(Y \le y) = P(2X + 5 \le y) = P(X \le \frac{y-5}{2}) = F_X(\frac{y-5}{2}) = \frac{y-5}{2}.$$

Differentiating $P(Y \le y)$ with respect to y, we get the probability density function of Y, for $5 \le y \le 7$,

$$f_Y(y) = \frac{1}{2}.$$

Problem 25. We have cdf of X,

$$F_X(x) = \int_0^x \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x}.$$

Now for $y \geq 0$, we have

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(X \leq \sqrt{y}) = 1 - e^{-\lambda\sqrt{y}}.$$

Differentiating $F_Y(y)$ with respect to y, we have

$$f_Y(y) = \frac{\lambda}{2} y^{-\frac{1}{2}} e^{-\lambda\sqrt{y}}.$$

Problem 26. **(a)** We first make the probability tables

X	0	2	3
prob.	0.3	0.1	0.6
Y	3	3	12

$\Rightarrow E(X) = 0 \cdot 0.3 + 2 \cdot 0.1 + 3 \cdot 0.6 = 2$

(b) $E(X^2) = 0{\cdot}0.3 + 4{\cdot}0.1 + 9{\cdot}0.6 = 5.8 \Rightarrow \text{Var}(X) = E(X^2) - E(X)^2 = 5.8 - 4 = 1.8.$

(c) $E(Y) = 3 \cdot 0.3 + 3 \cdot 0.1 + 12 \cdot 6 = 8.4.$

L(d) $F_Y(7) = P(Y \leq 7) = 0.4.$

Problem 27. **(a)** There are a number of ways to present this.
$X \sim 3\,\text{binomial}(25, 1/6)$, so

$$P(X = 3k) = \binom{25}{k}\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{25-k}, \quad \text{for } k = 0,\ 1,\ 2,\ \ldots,\ 25.$$

(b) $X \sim 3\,\text{binomial}(25, 1/6)$.

Recall that the mean and variance of binomial(n, p) are np and $np(1-p)$. So,

$$E(X) = 3\,E(textbinomial(25, 1/6)) = 3{\cdot}25/6 = 75/6, \text{ and } \text{Var}(X) = 9\,\text{Var}(textbinomial(25, 1/6)) = 9{\cdot}25($$

(c) $E(X+Y) = E(X) + E(Y) = 150/6 = 25.,\ E(2X) = 2E(X) = 150/6 = 25.$

$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) = 250/4.\ \text{Var}(2X) = 4\text{Var}(X) = 500/4.$

The means of $X + Y$ and $2X$ are the same, but $\text{Var}(2X) > \text{Var}(X+Y)$.

This makes sense because in $X+Y$ sometimes X and Y will be on opposite sides from the mean so distances to the mean will tend to cancel, However in $2X$ the distance to the mean is always doubled.

Problem 28. First we find the value of a:

$$\int_0^1 f(x)\,dx = 1 = \int_0^1 x + ax^2\,dx = \frac{1}{2} + \frac{a}{3} \;\Rightarrow\; a = 3/2.$$

The CDF is $F_X(x) = P(X \le x)$. We break this into cases:
(i) $b < 0 \;\Rightarrow\; F_X(b) = 0$.
(ii) $0 \le b \le 1 \;\Rightarrow\; F_X(b) = \displaystyle\int_0^b x + \frac{3}{2}x^2\,dx = \frac{b^2}{2} + \frac{b^3}{2}$.
(iii) $1 < x \;\Rightarrow\; F_X(b) = 1$.
Using F_X we get

$$P(.5 < X < 1) = F_X(1) - F_X(.5) = 1 - \left(\frac{.5^2 + .5^3}{2}\right) = \frac{13}{16}.$$

Problem 29. Let $Z = X + Y$. We'll build the probability table

Z	0	1	2
(X,Y)	(0,0)	(0,1), (1,0)	(1,1)
prob.	$(1-p)(1-q)$	$(1-p)q + p(1-q)$	pq
prob.	3/8	1/2	1/8

Not binomial (probabilites for binomial are 1/4, 1/2, 1/4).

Problem 30. **(a)** Note: $P(Y = 1) = P(X = 1) + P(X = -1)$.

Values a of Y:	0	1	4
PMF $p_Y(a)$:	1/8	3/8	1/2

(b) To distinguish the distribution functions we'll write F_x and F_Y.
Using the tables in part (a) and the definition $F_X(a) = P(X \le a)$ etc. we get

a:	1	3/4	$\pi - 3$
$F_X(a)$:	1/2	3/8	3/8
$F_Y(a)$:	1/2	1/8	1/8

Problem 31. The jumps in the distribution function are at 0, 1/2, 3/4. The value of $p(a)$ at at a jump is the height of the jump:

a:	0	1/2	3/4
$p(a)$:	1/3	1/6	1/2

Problem 32. (i) yes, discrete, (ii) no, (iii) no, (iv) no, (v) yes, continuous (vi) no (vii) yes, continuous, (viii) yes, continuous.

Problem 33. $P(1/2 \le X \le 3/4) = F(3/4) - F(1/2) = (3/4)^2 - (1/2)^2 = \boxed{5/16}$.

Problem 34. **(a)** $P(1/4 \le X \le 3/4) = F(3/4) - F(1/4) = \boxed{11/16 = .6875.}$

(b) $f(x) = F'(x) = 4x - 4x^3$ in [0,1].

Problem 35. We compute

$$P(X \geq 5) = 1 - P(X < 5) = 1 - \int_0^5 \lambda \mathrm{e}^{-\lambda x} dx = 1 - (1 - \mathrm{e}^{-5\lambda}) = \mathrm{e}^{-5\lambda}.$$

(b) We want $P(X \geq 15 | X \geq 10)$. First observe that $P(X \geq 15, X \geq 10) = P(X \geq 15)$. From similar computations in (a), we know

$$P(X \geq 15) = \mathrm{e}^{-15\lambda} \qquad P(X \geq 10) = \mathrm{e}^{-10\lambda}.$$

From the definition of conditional probability,

$$P(X \geq 15 | X \geq 10) = \frac{P(X \geq 15, X \geq 10)}{P(X \geq 10)} = \frac{P(X \geq 15)}{P(X \geq 10)} = \mathrm{e}^{-5\lambda}$$

Note: This is an illustration of the memorylessness property of the exponential distribution.

Problem 36. We have

$$F_X(x) = P(X \leq x) = P(3Z + 1 \leq x) = P(Z \leq \frac{x-1}{3}) = \Phi(\frac{x-1}{3}).$$

(b) Differentiating with respect to x, we have

$$f_X(x) = \frac{\mathrm{d}}{\mathrm{dx}} F_X(x) = \frac{1}{3}\phi(\frac{x-1}{3}).$$

Since $\phi(x) = (2\pi)^{-\frac{1}{2}} \mathrm{e}^{-\frac{x^2}{2}}$, we conclude

$$f_X(x) = \frac{1}{3\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-1)^2}{2 \cdot 3^2}},$$

which is the probability density function of the $N(1, 9)$ distribution. **Note:** The arguments in (a) and (b) give a proof that $3Z + 1$ is a normal random variable with mean 1 and variance 9. See Problem Set 3, Question 5.

(c) We have

$$P(-1 \leq X \leq 1) = P(-\frac{2}{3} \leq Z \leq 0) = \Phi(0) - \Phi(-\frac{2}{3}) \approx 0.2475$$

(d) Since $E(X) = 1$, $\mathrm{Var}(X) = 9$, we want $P(-2 \leq X \leq 4)$. We have

$$P(-2 \leq X \leq 4) = P(-3 \leq 3Z \leq 3) = P(-1 \leq Z \leq 1) \approx 0.68.$$

Problem 37. **(a)** Note, Y follows what is called a *log-normal distribution.*
$F_Y(a) = P(Y \le a) = P(e^Z \le a) = P(Z \le \ln(a)) = \boxed{\Phi(\ln(a)).}$
Differentiating using the chain rule:

$$f_y(a) = \frac{d}{da}F_Y(a) = \frac{d}{da}\Phi(\ln(a)) = \frac{1}{a}\phi(\ln(a)) = \boxed{\frac{1}{\sqrt{2\pi}\,a}\mathrm{e}^{-(\ln(a))^2/2}.}$$

(b) (i) We want to find $q_{.33}$ such that $P(Z \le q_{.33}) = .33$. That is, we want

$$\Phi(q_{.33}) = .33 \;\Leftrightarrow\; \boxed{q_{.33} = \Phi^{-1}(.33)}.$$

(ii) We want $q_{.9}$ such that

$$F_Y(q_{.9}) = .9 \;\Leftrightarrow\; \Phi(\ln(q_{.9})) = .9 \;\Leftrightarrow\; \boxed{q_{.9} = \mathrm{e}^{\Phi^{-1}(.9)}}.$$

(iii) As in (ii) $q_{.5} = \mathrm{e}^{\Phi^{-1}(.5)} = \mathrm{e}^0 = \boxed{1}$.

Problem 38. **(a)** **answer:** $\mathrm{Var}(X_j) = 1 = E(X_j^2) - E(X_j)^2 = E(X_j^2)$. QED

(b) $E(X_j^4) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{\infty} x^4\mathrm{e}^{-x^2/2}\,dx.$

(Extra credit) By parts: let $u = x^3$, $v' = x\mathrm{e}^{-x^2/2} \;\Rightarrow\; u' = 3x^2$, $v = -\mathrm{e}^{-x^2/2}$

$$E(X_j^4) = \frac{1}{\sqrt{2\pi}}\left[x^3\mathrm{e}^{-x^2/2}\Big|_{infty}^{\infty} + \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} 3x^2\mathrm{e}^{-x^2/2}\,dx\right]$$

The first term is 0 and the second term is the formula for $3E(X_j^2) = 3$ (by part (a)). Thus, $E(X_j^4) = 3$.

(c) **answer:** $\mathrm{Var}(X_j^2) = E(X_j^4) - E(X_j^2)^2 = 3 - 1 = 2$. QED

(d) $E(Y_{100}) = E(100X_j^2) = 100$. $\mathrm{Var}(Y_{100}) = 100\mathrm{Var}(X_j) = 200$.

The CLT says Y_{100} is approximately normal. Standardizing gives

$$P(Y_{100} > 110) = P\left(\frac{Y_{100} - 100}{\sqrt{200}}) > \frac{10}{\sqrt{200}}\right) \approx P(Z > 1/\sqrt{2}) = \boxed{.24}.$$

This last value was computed using `1 - pnorm(1/sqrt(2),0,1)`.

Problem 39.

(a) We did this in class. Let $\phi(z)$ and $\Phi(z)$ be the PDF and CDF of Z.
$F_Y(y) = P(Y \le y) = P(aZ + b \le y) = P(Z \le (y - b)/a) = \Phi((y - b)/a)$.
Differentiating:

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}\Phi((y - b)/a) = \frac{1}{a}\phi((y - b)/a) = \frac{1}{\sqrt{2\pi}\,a}\,\mathrm{e}^{-(y-b)^2/2a^2}.$$

Since this is the density for $\mathrm{N}(b, a^2)$ we have shown $Y \sim \mathrm{N}(b, a^2)$.

(b) By part (a), $Y \sim \mathrm{N}(\mu, \sigma^2) \Rightarrow Y = \sigma Z + \mu$.
But, this implies $(Y - \mu)/\sigma = Z \sim \mathrm{N}(0, 1)$. QED

Problem 40. a) $E(W) = 3E(X) - 2E(Y) + 1 = 6 - 10 + 1 = \boxed{-3}$
$\mathrm{Var}(W) = 9\mathrm{Var}(X) + 4\mathrm{Var}(Y) = 45 + 36 = \boxed{81}$
b) Since the sum of independent normal is normal part (a) shows: $W \sim N(-3, 81)$.
Let $Z \sim N(0, 1)$. We *standardize* W: $P(W \le 6) = P\left(\frac{W+3}{9} \le \frac{9}{9}\right) = P(Z \le 1) \approx \boxed{.84.}$

Problem 41. Let $X \sim \mathrm{U}(a, b)$. Compute $E(X)$ and $\mathrm{Var}(X)$.
Method 1
$U(a, b)$ has density $f(x) = \frac{1}{b-a}$ on $[a, b]$. So,

$$E(X) = \int_a^b x f(x)\, dx = \frac{1}{b-a}\int_a^b x\, dx = \left.\frac{x^2}{2(b-a)}\right|_a^b = \frac{b^2 - a^2}{2(b-a)} = \boxed{\frac{a+b}{2}}.$$

$$E(X^2) = \int_a^b x^2 f(x)\, dx = \frac{1}{b-a}\int_a^b x^2\, dx = \left.\frac{x^3}{3(b-a)}\right|_a^b = \frac{b^3 - a^3}{3(b-a)}.$$

Finding $\mathrm{Var}(X)$ now requires a little algebra,

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = \frac{b^3 - a^3}{3(b-a)} - \frac{(b+a)^2}{4}$$

$$= \frac{4(b^3 - a^3) - 3(b-a)(b+a)^2}{12(b-a)} = \frac{b^3 - 3ab^2 + 3a^2 b - a^3}{12(b-a)} = \frac{(b-a)^3}{12(b-a)} = \boxed{\frac{(b-a)^2}{12}}.$$

Method 2
There is an easier way to find $E(X)$ and $\mathrm{Var}(X)$.
Let $U \sim \mathrm{U}(a, b)$. Then the calculations above show $E(U) = 1/2$ and $(E(U^2) = 1/3$ $\Rightarrow$ $\mathrm{Var}(U) = 1/3 - 1/4 = 1/12$.
Now, we know $X = (b-a)U + a$, so $E(X) = (b-a)E(U) + a = (b-a)/2 + a = (b+a)/2$ and $\mathrm{Var}(X) = (b-a)^2\mathrm{Var}(U) = (b-a)^2/12$.

Problem 42. In $n + m$ independent Bernoulli(p) trials, let S_n be the number of successes in the first n trials and T_m the number of successes in the last m trials.

(a) **(a)** $S_n \sim$ Binomial(n, p), since it is the number of successes in n independent Bernoulli trials.

(b) **(b)** $T_m \sim$ Binomial(m, p), since it is the number of successes in m independent Bernoulli trials.

(c) (c) $S_n + T_m \sim$ Binomial$(n+m,\, p)$, since it is the number of successes in $n+m$ independent Bernoulli trials.

(d) **(d)** Yes, S_n and T_m are independent. We haven't given a formal definition of independent random variables yet. But, we know it means that knowing S_n gives no information about T_m. This is clear since the first n trials are independent of the last m.

Problem 43. Compute the median for the exponential distribution with parameter λ. The density for this distribution is $f(x) = \lambda\, e^{-\lambda x}$. We know (or can compute) that the distribution function is $F(a) = 1 - e^{-\lambda a}$. The median is the value of a such that $F(a) = .5$. Thus, $1 - e^{-\lambda a} = 0.5 \;\Rightarrow\; 0.5 = e^{-\lambda a} \;\Rightarrow\; \log(0.5) = -\lambda a \;\Rightarrow$ $\boxed{a = \log(2)/\lambda.}$

Problem 44. Let X = the number of heads on the first 2 flips and Y the number in the last 2. Considering all 8 possibe tosses: HHH, HHT etc we get the following joint pmf for X and Y

Y/X	0	1	2	
0	1/8	1/8	0	1/4
1	1/8	1/4	1/8	1/2
2	0	1/8	1/8	1/4
	1/4	1/2	1/4	1

Using the table we find

$$E(XY) = \frac{1}{4} + 2\frac{1}{8} + 2\frac{1}{8} + 4\frac{1}{8} = \frac{5}{4}.$$

We know $E(X) = 1 = E(Y)$ so

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{5}{4} - 1 = \frac{1}{4}.$$

Since X is the sum of 2 independent Bernoulli(.5) we have $\sigma_X = \sqrt{2/4}$

$$\text{Cor}(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{1/4}{(2)/4} = \frac{1}{2}.$$

Problem 45. As usual let X_i = the number of heads on the i^{th} flip, i.e. 0 or 1.

Let $X = X_1 + X_2 + X_3$ the sum of the first 3 flips and $Y = X_3 + X_4 + X_5$ the sum of the last 3. Using the algebraic properties of covariance we have

$$\begin{aligned}
\text{Cov}(X,Y) &= \text{Cov}(X_1 + X_2 + X_3, X_3 + X_4 + X_5) \\
&= \text{Cov}(X_1, X_3) + \text{Cov}(X_1, X_4) + \text{Cov}(X_1, X_5) \\
&+ \text{Cov}(X_2, X_3) + \text{Cov}(X_2, X_4) + \text{Cov}(X_2, X_5) \\
&+ \text{Cov}(X_3, X_3) + \text{Cov}(X_3, X_4) + \text{Cov}(X_3, X_5)
\end{aligned}$$

Because the X_i are independent the only non-zero term in the above sum is $\text{Cov}(X_3 X_3) = \text{Var}(X_3) = \dfrac{1}{4}$
Therefore, $\text{Cov}(X,Y) = \frac{1}{4}$.
We get the correlation by dividing by the standard deviations. Since X is the sum of 3 independent Bernoulli(.5) we have $\sigma_X = \sqrt{3/4}$

$$\text{Cor}(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{1/4}{(3)/4} = \frac{1}{3}.$$

Problem 46.

answer: $U = X + Y$ takes values 0, 1, 2 and $V = |X - Y|$ takes values 0, 1.
The table is computed as follows:
$P(U = 0, V = 0) = P(X = 0, Y = 0) = 1/4,$
$P(U = 1, V = 0) = 0,$
$P(U = 2, V = 0) = P(X = 1, Y = 1) = 1/4.$
$P(U = 0, V = 1) = 0,$
$P(U = 1, V = 1) = P(X = 1, Y = 0) + P(X = 0, Y = 1) = 1/2,$
$P(U = 2, V = 1) = 0.$

$V \backslash U$	0	1	2	P_V
0	1/4	0	1/4	1/2
1	0	1/2	0	1/2
P_U	1/4	1/2	1/4	1

Problem 47. **(a)** The joint distribution is given by

$Y \backslash X$	1	2	3	
1	$\frac{1168}{5383}$	$\frac{825}{5383}$	$\frac{305}{5383}$	$\frac{2298}{5383}$
2	$\frac{573}{5383}$	$\frac{1312}{5383}$	$\frac{1200}{5383}$	$\frac{3085}{5383}$
	$\frac{1741}{5383}$	$\frac{2137}{5383}$	$\frac{1505}{5383}$	1

with the marginal distribution of X at right and of Y at bottom.
(b) X and Y are dependent because, for example,

$$P(X = 1 \text{ and } Y = 1) = \frac{1168}{5383}$$

is not equal to

$$P(X = 1)P(Y = 1) = \frac{1741}{5383} \cdot \frac{2298}{5383}.$$

Problem 48. **(a)** Here we have two continuous random variables X and Y with going potability density function

$$f(x,y) = \frac{12}{5} xy(1 + y) \quad \text{for } 0 \le x \le 1 \text{ and } 0 \le y \le 1,$$

and $f(x,y) = 0$ otherwise. So

$$P(\frac{1}{4} \le X \le \frac{1}{2}, \frac{1}{3} \le Y \le \frac{2}{3}) = \int_{\frac{1}{4}}^{\frac{1}{2}} \int_{\frac{1}{3}}^{\frac{2}{3}} f(x,y)dy\,dx = \frac{41}{720}.$$

(b) $F(a,b) = \int_0^a \int_0^b f(x,y)dy\,dx = \frac{3}{5}a^2b^2 + \frac{2}{5}a^2b^3$ for $0 \le a \le 1$ and $0 \le b \le 1$.
(c) Since $f(x,y) = 0$ for $y > 1$, we have

$$F_X(a) = \lim_{b\to\infty} F(a,b) = F(a,1) = a^2.$$

(d) For $0 \le x \le 1$, we have

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^1 f(x,y)dy = 2x.$$

This is consistent with (c) because $\frac{d}{dx}(x^2) = 2x$.
(e) We first compute $f_Y(y)$ for $0 \le y \le 1$ as

$$f_Y(y) = \int_0^1 f(x,y)dx = \frac{6}{5}y(y+1).$$

Since $f(x,y) = f_X(x)f_Y(y)$, we conclude that X and Y are independent.

Problem 49. **(a)** First note $E(X+s) = E(X)+s$, thus $X+s-E(X+s) = X-E(X)$.
Likewise $Y+u-E(Y+u) = Y-E(Y)$.
Now using the definition of covariance we get

$$\begin{aligned}\text{Cov}(X+s, Y+u) &= E((X+s-E(X+s)) \cdot (Y+u-E(Y+u))) \\ &= E((X-E(X)) \cdot (Y-E(Y))) \\ &= \text{Cov}(X,Y).\end{aligned}$$

(b) For practice, here we'll use the formula (see problem 11)
$\text{Cov}(X,Y) = E(XY) - E(X)E(Y)$.

$$\begin{aligned}\text{Cov}(rX, tY) &= E((rX)(tY)) - E(rX)E(tY) \\ &= rt(E(XY) - E(X)E(Y)) \\ &= rt\text{Cov}(X,Y)\end{aligned}$$

(c) We have

$$\text{Cov}(rX+s, tY+u) = \text{Cov}(rX, tY) = rt\text{Cov}(X,Y)$$

where the first equality is by (a) and the second equality is by (b).

Problem 50. Using linearity of expectation, we have

$$\begin{aligned}\text{Cov}(X, Y) &= E((X - E(X))(Y - E(Y))) \\ &= E(XY - E(X)Y - E(Y)X + E(X)E(Y)) \\ &= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y).\end{aligned}$$

Problem 51. **(a)** The marginal probability $P_Y(1) = 1/2$
$\Rightarrow\ P(X = 0, Y = 1) = P(X = 2, Y = 1) = 0.$
Now each column has one empty entry. This can be computed by making the column add up to the given marginal probability.

$Y \backslash X$	0	1	2	P_Y
-1	1/6	1/6	1/6	1/2
1	0	1/2	0	1/2
P_X	1/6	2/3	1/6	1

(b) No, X and Y are not independent.
For example, $P(X = 0, Y = 1) = 0 = 1/12 = P(X = 0) \cdot P(Y = 1)$.

Problem 52. For shorthand, let $P(X = a, Y = b) = p(a, b)$.
(a) $P(X = Y) = p(1, 1) + p(2, 2) + p(3, 3) + p(4, 4) = \boxed{34/136.}$
(b) $P(X + Y = 5) = p(1, 4) + p(2, 3) + p(3, 2) + p(4, 1) = \boxed{34/136.}$
(c) $P(1 < X \leq 3, 1 < Y \leq 3) =$ sum of middle 4 probabilities in table $= \boxed{34/136.}$
(d) $\{1, 4\} \times \{1, 4\} = \{(1, 1), (1, 4), (4, 1), (4, 4)\ \Rightarrow$ prob. $= \boxed{34/136.}$

Problem 53. **(a)** X and Y are independent, so the table is computed from the product of the known marginal probabilities. Since they are independent, $\text{Cov}(X, Y) = 0$.

$Y \backslash X$	0	1	P_Y
0	1/8	1/8	1/4
1	1/4	1/4	1/2
2	1/8	1/8	1/4
P_X	1/2	1/2	1

(b) The sample space is $\Omega = \{$HHH, HHT, HTH, HTT, THH, THT, TTH, TTT$\}$.
$P(X = 0, Z = 0) = P(\{TTH, TTT\}) = 1/4.$
$P(X = 0, Z = 1) = P(\{THH, THT\}) = 1/4.$
$P(X = 0, Z = 2) = 0.$
$P(X = 1, Z = 0) = 0.$
$P(X = 1, Z = 1) = P(\{HTH, HTT\}) = 1/4.$
$P(X = 1, Z = 2) = P(\{HHH, HHT\}) = 1/4.$

$Z \backslash X$	0	1	P_Z
0	1/4	0	1/4
1	1/4	1/4	1/2
2	0	1/4	1/4
P_X	1/2	1/2	1

$\text{Cov}(X, Z) = E(XZ) - E(X)E(Z)$.
$E(X) = 1/2, \quad E(Z) = 1,\ E(XZ) = \sum x_i y_j p(x_i, y_j) = 3/4$.
$\Rightarrow\ \text{Cov}(X, Z) = 3/4 - 1/2 = \boxed{1/4.}$

Problem 54. **(a)**

Y \ X	-2	-1	0	1	2	
0	0	0	1/5	0	0	1/5
1	0	1/5	0	1/5	0	2/5
4	1/5	0	0	0	1/5	2/5
	1/5	1/5	1/5	1/5	1/5	1

Each column has only one nonzero value. For example, when $X = -2$ then $Y = 4$, so in the $X = -2$ column, only $P(X = -2, Y = 4)$ is not 0.

(b) Using the marginal distributions: $E(X) = \frac{1}{5}(-2 - 1 + 0 + 1 + 2) = 0$.
$$E(Y) = 0 \cdot \frac{1}{5} + 1 \cdot \frac{2}{5} + 4 \cdot \frac{2}{5} = 2.$$

(c) We show the probabilities don't multiply:
$P(X = -2,\ Y = 0) = 0 = P(X = -2) \cdot P(Y = 0) = 1/25$.
Since these are not equal X and Y are not independent. (It is obvious that X^2 is not independent of X.)

(d) Using the table from part (a) and the means computed in part (d) we get:
$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{1}{5}(-2)(4) + \frac{1}{5}(-1)(1) + \frac{1}{5}(0)(0) + \frac{1}{5}(1)(1) + \frac{1}{5}(2)(4) = 0.$$

Problem 55. **(a)** $F(a, b) = P(X \le a, Y \le b) = \int_0^a \int_0^b (x + y)\, dy\, dx.$

Inner integral: $xy + \left.\frac{y^2}{2}\right|_0^b = xb + \frac{b^2}{2}$. Outer integral: $\left.\frac{x^2}{2}b + \frac{b^2}{2}x\right|_0^a = \frac{a^2 b + ab^2}{2}$.

So $\boxed{F(x, y) = \frac{x^2 y + xy^2}{2}}$ and $\boxed{F(1, 1) = 1.}$

(b) $f_X(x) = \int_0^1 f(x, y)\, dy = \int_0^1 (x + y)\, dy = xy + \left.\frac{y^2}{2}\right|_0^1 = \boxed{x + \frac{1}{2}.}$

By symmetry, $\boxed{f_Y(y) = y + 1/2.}$

(c) To see if they are independent we check if the joint density is the product of the marginal densities.
$f(x, y) = x + y, \quad f_X(x) \cdot f_Y(y) = (x + 1/2)(y + 1/2)$.
Since these are not equal, $\boxed{X \text{ and } Y \text{ are not independent.}}$

(d) $E(X) = \int_0^1 \int_0^1 x(x+y)\,dy\,dx = \int_0^1 \left[x^2 y + x\frac{y^2}{2}\Big|_0^1 \right] dx = \int_0^1 x^2 + \frac{x}{2}\,dx = \boxed{\frac{7}{12}.}$

(Or, using (b), $E(X) = \int_0^1 x f_X(x)\,dx = \int_0^1 x(x+1/2)\,dx = 7/12$.)

By symmetry $\boxed{E(Y) = 7/12.}$

$E(X^2+Y^2) = \int_0^1 \int_0^1 (x^2+y^2)(x+y)\,dy\,dx = \boxed{\frac{5}{6}.}$

$E(XY) = \int_0^1 \int_0^1 xy(x+y)\,dy\,dx = \boxed{\frac{1}{3}.}$

$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{3} - \frac{49}{144} = \boxed{-\frac{1}{144}.}$

Problem 56.
Standardize:

$$\begin{aligned} P\left(\sum_i X_i < 30\right) &= P\left(\frac{\frac{1}{n}\sum X_i - \mu}{\sigma/\sqrt{n}} < \frac{30/n - \mu}{\sigma/\sqrt{n}}\right) \\ &\approx P\left(Z < \frac{30/100 - 1/5}{1/30}\right) \quad \text{(by the central limit theorem)} \\ &= P(Z < 3) \\ &= 1 - .0013 = .9987 \ \text{(from the table)} \end{aligned}$$

Problem 57. If $p < .5$ your expected winnings on any bet is negative, if $p = .5$ it is 0, and if $p > .5$ is is positive. By making a lot of bets the minimum strategy will 'win' you close to the expected average. So if $p \leq .5$ you should use the maximum strategy and if $p > .5$ you should use the minumum strategy.

Problem 58. Let $\overline{X} = \frac{X_1 + \ldots + X_{144}}{144} \Rightarrow E(\overline{X}) = 2$, and $\sigma_{\overline{X}} = 2/12 = 1/6$.
($\sqrt{n} = 12$)

A chain of algebra gives $P(X_1 + \ldots + X_{144}) + P\left(\overline{X} > \frac{264}{144}\right) = P\left(\overline{X} > 1.8333\right)$.

Standardization gives $P(\overline{X} > 1.8333) = P\left(\frac{\overline{X}-2}{1/6} > \frac{1.8333-2}{1/6}\right) = P\left(\frac{\overline{X}-2}{1/6} > -1.0\right)$

Now, the Central limit theorem says $P\left(\frac{\overline{X}-2}{1/6} > -1.0\right) \approx P(Z > -1) = \boxed{.84}$

Problem 59. This is an application of the Central Limit Theorem.
$X \sim \text{Bin}(n,p)$ means X is the sum of n i.i.d. Bernoulli(p) random variables.

We know $E(X) = np$ and $\mathrm{Var}(X) = np(1-p)$, $\sigma_X = \sqrt{np(1-p)}$.
Since X is a sum of i.i.d. random variables, the CLT theorem says $X \approx \mathrm{N}(np, np(1-p))$.
Standardization then gives $\dfrac{X - np|}{\sqrt{np(1-p)}} \approx \mathrm{N}(0,1)$.

Problem 60. **(a)** When this question was asked in a study, the number of undergraduates who chose each option was 21, 21, and 55, respectively. This shows a lack of intuition for the relevance of sample size on deviation from the true mean (i.e., variance).

(b) The random variable X_L, giving the number of boys born in the larger hospital on day i, is governed by a Bin(45, .5) distribution. So L_i has a Ber(p_L) distribution with

$$p_L = P(X > 27) = \sum_{k=28}^{45} \left(\begin{array}{c} 45 \\ k \end{array} \right) .5^{45} \approx .068$$

Similarly, the random variable X_S, giving the number of boys born in the smaller hospital on day i, is governed by a Bin(15, .5) distribution. So S_i has a Ber(p_S) distribution with

$$p_S = P(X_S > 9) = \sum_{k=10}^{15} \left(\begin{array}{c} 15 \\ k \end{array} \right) .5^{15} \approx .151$$

We see that p_S is indeed greater than p_L, consistent with (ii).

(c) Note that $L = \sum_{i=1}^{365} L_i$ and $S = \sum_{i=1}^{365} S_i$. So L has a Bin($365, p_L$) distribution and S has a Bin($365, p_S$) distribution. Thus

$$\begin{aligned} E(L) &= 365p_L \approx 25 \\ E(S) &= 365p_S \approx 55 \\ \mathrm{Var}(L) &= 365p_L(1-p_L) \approx 23 \\ \mathrm{Var}(S) &= 365p_S(1-p_S) \approx 47 \end{aligned}$$

(d) mean + sd in each case:

For L, $q_{.84} \approx 25 + \sqrt{23}$.

For S, $q_{.84} \approx 55 + \sqrt{47}$.

(e) Since L and S are independent, their joint distribution is determined by multiplying their individual distributions. Both L and S are binomial with $n = 365$ and p_L and p_S computed above. Thus

$$p_{l,s}P(L = i \text{ and } S = j) = p(i,j) = \binom{365}{i} p_L^i (1-p_L)^{365-i} \binom{365}{j} p_S^j (1-p_S)^{365-j}$$

Thus

$$P(L > S) = \sum_{i=0}^{364} \sum_{j=i+1}^{365} p(i,j) \approx .0000916$$

(We used R to do the computations.)