6.431 Final exam solutions

Problem 1: *(50 points)*

Everyone is invited to the annual Probability Party at M.I.T. The doors open at 7pm. Everyone who arrives before 7pm is sent away. Everyone who arrives after 7pm stays at least until midnight. Undergraduate students arrive to the party according to a Poisson process with rate λ_u students per hour. Graduate students arrive according to a Poisson process with rate λ_g students per hour, independently of the undergraduate students. Independently of everything else, each student arriving at the party is a female with probability p.

a. *(5 points)* Let A be the event that there are exactly 4 students at the party at 9pm. What is the probability of event A?

$\mathbf{P}(A) = \frac{(2(\lambda_u+\lambda_g))^4}{4!} e^{-2(\lambda_u+\lambda_g)}$

The combined arrival process is also a Poisson process with rate $\lambda_u + \lambda_g$. The total number of arrivals in 2 hours is a Poisson random variable with parameter $2(\lambda_u + \lambda_g)$.

b. *(5 points)* Let B be the event that out of the first 10 students in the party, exactly 4 are graduate students. What is the probability of event B?

$\mathbf{P}(B) = \binom{10}{4} \left(\frac{\lambda_g}{\lambda_u+\lambda_g}\right)^4 \left(\frac{\lambda_u}{\lambda_u+\lambda_g}\right)^6$

Each arrival of the combined process is a graduate student with probability $\frac{\lambda_g}{\lambda_u+\lambda_g}$ or an undergraduate student with probability $\frac{\lambda_u}{\lambda_u+\lambda_g}$. The number of graduate student in the first 10 arrivals is a binomial random variable, and the probability of event B is the probability that this binomial random variable is equal to 4.

c. *(6 points)* Let W be the time of arrival of the 7th woman to the party. Find the expectation, the variance and the transform of W.

$\mathbf{E}[W] = \frac{7}{p(\lambda_u+\lambda_g)}$

$var(W) = \frac{7}{p^2(\lambda_u+\lambda_g)^2}$

$M_W(s) = \left(\frac{p(\lambda_u+\lambda_g)}{p(\lambda_u+\lambda_g)-s}\right)^7 \qquad s < p(\lambda_u + \lambda_g)$

The process of arrivals of women to this party is a also a Poisson process with rate $p(\lambda_u+\lambda_g)$. W is therefore an Erlang random variable of order 7. Another way to think about it is to realize that $W = T_1 + T_2 + \cdots + T_7$ where T_i's are independent of each other and each T_i is an exponential random variable with parameter $p(\lambda_u + \lambda_g)$, leading to the following answers:
$\mathbf{E}[W] = 7\mathbf{E}[T] = \frac{7}{p(\lambda_u+\lambda_g)}$

$var(W) = 7var(T) = \frac{7}{p^2(\lambda_u+\lambda_g)^2}$

$M_W(s) = (M_T(s))^7 = \left(\frac{p(\lambda_u+\lambda_g)}{p(\lambda_u+\lambda_g)-s}\right)^7 \qquad s < p(\lambda_u+\lambda_g)$

d. *(6 points)* Let C be the event that the 3rd undergraduate student arrives before the 2nd graduate student. What is the probability of event C?

$$\begin{aligned}\mathbf{P}(C) &= \left(\frac{\lambda_u}{\lambda_u+\lambda_g}\right)^3 + 3\frac{\lambda_g}{\lambda_u+\lambda_g}\left(\frac{\lambda_u}{\lambda_u+\lambda_g}\right)^3 \\ &= \left(\frac{\lambda_u}{\lambda_u+\lambda_g}\right)^4 + 4\frac{\lambda_g}{\lambda_u+\lambda_g}\left(\frac{\lambda_u}{\lambda_u+\lambda_g}\right)^3 \\ &= \frac{\lambda_u^4+4\lambda_u^3\lambda_g}{(\lambda_u+\lambda_g)^4}\end{aligned}$$

<u>Solution A</u>

We need to add up the probability of all possible arrival sequences that map to this event. There are 4 such sequences, $\{uuu\}, \{guuu\}, \{uguu\}, \{uugu\}$. Since these event are mutually exclusive, we can simply add the probabilities:

$$\mathbf{P}(C) = \left(\frac{\lambda_u}{\lambda_u+\lambda_g}\right)^3 + 3\frac{\lambda_g}{\lambda_u+\lambda_g}\left(\frac{\lambda_u}{\lambda_u+\lambda_g}\right)^3$$

<u>Solution B</u>

Looking at the first four arrivals, we are interested in sequences with either 4 undergraduates, or 3 undergraduate and 1 graduate students. The probability of having 4 undergraduate students is $\left(\frac{\lambda_u}{\lambda_u+\lambda_g}\right)^4$ and the probability of the second type is $4\frac{\lambda_g}{\lambda_u+\lambda_g}\left(\frac{\lambda_u}{\lambda_u+\lambda_g}\right)^3$. Adding these up, yield

$$\mathbf{P}(C) = \left(\frac{\lambda_u}{\lambda_u+\lambda_g}\right)^4 + 4\frac{\lambda_g}{\lambda_u+\lambda_g}\left(\frac{\lambda_u}{\lambda_u+\lambda_g}\right)^3$$

It is easy to check that both solutions simplify to the same answer above.

e. *(7 points)* At 1am no one else is allowed into the party, and there are only 5 remaining guests. All guests depart independently of each other. Each guest's departure time is an exponential random variable with rate μ. What is the expected time until the first of these guests leaves? What is the expected time until the last of the guests leaves?

$\mathbf{E}[\text{Time until first departure}] = \frac{1}{5\mu}$

$\mathbf{E}[\text{Time until last departure}] = \frac{1}{5\mu} + \frac{1}{4\mu} + \frac{1}{3\mu} + \frac{1}{2\mu} + \frac{1}{\mu}$

In the first part, the quantity we are trying to calculate is the expectation of the time until the first arrival of five competing exponentials, each with parameter μ. Each of these exponential random variables corresponds to the time until the first arrival of an imaginary Poisson processes with rate μ. Consider the processes created by merging these five processes. The resulting process is a Poisson process with rate 5μ. Each arrival of this merged process corresponds to a departure of one of the guests from the party. So the expected time until the first arrival of such process corresponds to the expected time until the departure of the first guest. Since the process has rate 5μ, this expectation is $\frac{1}{5\mu}$.

Using a similar approach, after the first departure, we are left with only four competing exponentials. Based on the memoryless property of exponentials and using a similar argument, the expected time until the second departure is $\frac{1}{4\mu}$. Continuing a similar argument, the time until the last departure is $\frac{1}{5\mu}+\frac{1}{4\mu}+\frac{1}{3\mu}+\frac{1}{2\mu}+\frac{1}{\mu}$.

f. *(7 points)* Alice leaves the party at some random time after midnight and goes to the M.I.T. bus stop to catch the ride home. Her route is served by one bus that runs 24 hours a day. The time between any two consecutive times when the bus arrives at the M.I.T. stop is distributed uniformly between 40min and 60min, independently of what happens on all other runs. Let L be the time from the last bus arrival at the M.I.T bus stop to the time when Alice gets to the bus stop. Find the expectation of L.

$\mathbf{E}[L] = \frac{\mathbf{E}[X^2]}{2\mathbf{E}[X]} = \frac{var(x)+\mathbf{E}[X]^2}{2\mathbf{E}[X]} = \frac{\frac{20^2}{12}+50^2}{2\cdot 50} = 25\frac{1}{3}$

First, let's find the expectation and variance of the inter-arrival time X:

$\mathbf{E}[X] = \frac{60+40}{2} = 50$
$var(X) = \frac{(60-40)^2}{12} = \frac{400}{12}$

We need to find the conditional distribution of the length of inter-arrival time, T, conditioned on the event that Alice arrived during that inter-arrival time. We know that this probability should be proportional to the length of the inter-arrival time, i.e.

$f_T(x) = \alpha x f_X(x) \qquad 40 \le x \le 60$

Furthermore, $f_T(x)$ must integrate to 1. So $\alpha = \frac{1}{\mathbf{E}[X]}$.

$f_T(x) = \frac{1}{\mathbf{E}[X]} x f_X(x) \qquad 40 \le x \le 60$

Therefore,

$\mathbf{E}[T] = \int \frac{1}{\mathbf{E}[X]} x^2 f_X(x)dx = \frac{\mathbf{E}[X^2]}{\mathbf{E}[X]}$.

By interated expectation, $\mathbf{E}[L] = \mathbf{E}[\mathbf{E}[L|T]] = \mathbf{E}[T/2]$.

Using the expectation and the variance for X, we get the desired result.

g. *(7 points)* Student ages are distributed exponentially with parameter λ_a. The gate keeper at the party asks every guest how old they are, but the students always exaggerate their age. Independently of his/her true age, any particular student adds an amount that is exponentially distributed with parameter λ_d when answering the gate keeper's questions. Let the true age of a particular student be A, and the age this student told the gate keeper be X. Find the *linear* least squares estimator of A from X.

$$\hat{A}_{LLSE}(X) = \frac{1}{\lambda_a} + \frac{\lambda_d^2}{\lambda_a^2 + \lambda_d^2}(X - \frac{1}{\lambda_a} - \frac{1}{\lambda_d}) = \frac{\lambda_d^2}{\lambda_a^2 + \lambda_d^2}X + \frac{\lambda_a - \lambda_d}{\lambda_a^2 + \lambda_d^2}$$

We note that $X = A + D$, where X and A are as defined above and D is an exponential random variable with parameter λ_d. D is independent of A. Based on the LLSE formulation, we know:

$\hat{A}_{LLSE}(X) = \mathbf{E}[A] + \frac{cov(A,X)}{var(X)}(X - \mathbf{E}[X])$

$\mathbf{E}[A] = \frac{1}{\lambda_a}$

$\mathbf{E}[X] = \mathbf{E}[A] + \mathbf{E}[D] = \frac{1}{\lambda_a} + \frac{1}{\lambda_d}$

$var(X) = var(A) + var(D) = \frac{1}{\lambda_a^2} + \frac{1}{\lambda_d^2}$

where we used the fact that A and D are independent random variables.

$Cov(X, A) = Cov(A + D, A) = Cov(D, A) + Cov(A, A) = \frac{1}{\lambda_a^2}$

where we used $X = A + D$ and the fact D and A are independent and, hence, their covariance is 0.

Substituting all these quantities in the LLSE formula, we get the desired result.

h. *(7 points)* In this question, assume $\lambda_a = \lambda_d$. Find the *best* least squares estimator of A from X. Explain how it compares with your answer in the previous question.

$\hat{A}_{LSE}(X) = \frac{X}{2}$

Compare with your answer from (g), explain:

In this case, the LSE is linear which implies that the optimal LLSE must also be the LSE. Indeed, substituting $\lambda_a = \lambda_d$ in the answer for part (g), we get $\frac{X}{2}$. The two estimators are identical.

The optimal least square estimator is given by:

$\hat{A}_{LSE}(X) = \mathbf{E}[A|X]$

To calculate this conditional expectation, we need to find the conditional PDF of A given X. Using Bayes' rule, we have

$$f_{A|X}(a|x) = \frac{f_{X,A}(x,a)}{f_X(x)} = \frac{f_{X|A}(x|a)f_A(a)}{f_X(x)}$$

We note that X is the sum of two exponentials with parameter λ_a, i.e. it has an Erlang, order 2, distribution. Furthermore, both f_A and $f_{X|A}$ are exponential distributions with parameter λ_a. Substituting the densities into the above relation, we get:

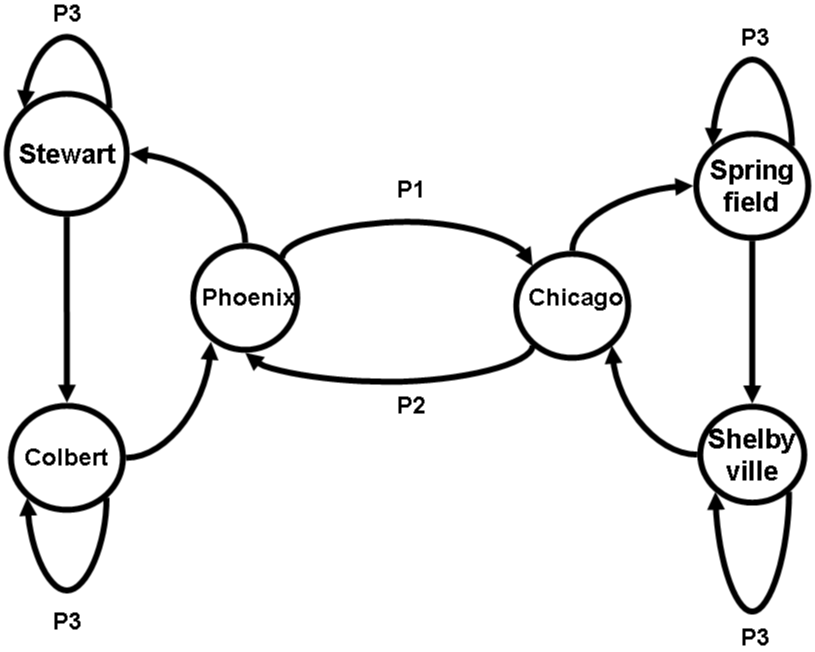
$$f_{A|X}(a|x) = \frac{\lambda_a e^{-\lambda_a(x-a)}\lambda_a e^{-\lambda_a a}}{\lambda_a^2 e^{-\lambda_a x} x} = \frac{1}{x}, \qquad 0 \le a \le x$$

Thus, conditioned on X, A is uniformly distributed between 0 and X, and $\mathbf{E}[A|X] = \frac{X}{2}$.

Problem 2: *(50 points)*

Bernoulli Airline services six airports with a single airplane. The plane makes exactly one flight every day. All flights take place in the middle of the day. The plane schedule can be described by a Markov process shown in the figure below, where p_1, p_2 and p_3 denote transition probabilities. The plane flies among two large cities, Phoenix and Chicago, and four small towns. When the plane is in Phoenix, sometimes it will fly back to Chicago but otherwise it will fly from Phoenix to the small town of Stewart, and then from Stewart to the small town of Colbert, and then finally back to Phoenix. When the plane is in Chicago, sometimes it will fly back to Phoenix, but otherwise it will fly from Chicago to the small town of Springfield, and then from Springfield it will fly to the small town of Shelbyville, and then back to Chicago.

The plane can fly from the large cities regardless of the weather, but it cannot take off from the small airports when the weather is bad. If the weather is bad and the plane is in the small town, it is stuck in that town until the weather is nice again and it can take off. The chance of bad weather on any particular day is independent of what happened on any other day and is equal to p_3.



a. *(5 points)* In this question, assume that the plane starts in Phoenix on Monday morning. Let A be the event that it will be in Stewart on Thursday morning. Find the probability of A.

$\mathbf{P}(A) = p_1 p_2 (1 - p_1) + (1 - p_1) p_3^2$

There are two possible paths that satisfy event A. These are:

Phoenix → Chicago → Phoenix → Stewart

Phoenix → Stewart → Stewart → Stewart

These outcomes are mutually exclusive, thus their probabilities can be added. The probability of the first outcome is $p_1p_2(1-p_1)$. The probability of the second outcome is $(1-p_1)p_3^2$. Thus, $\mathbf{P}(A) = p_1p_2(1-p_1) + (1-p_1)p_3^2$

b. *(5 points)* Suppose the plane is now in the town of Stewart. Let N be the number of days it will take for the plane to arrive in Colbert for the first time from now. Determine the PMF of N.

$$p_N(n) = p_3^{n-1}(1-p_3), \quad n = 1, 2, ...$$

We note from the Markov chain that N is a geometric random variable with probability of success $1 - p_3$.

c. *(6 points)* Suppose the plane is now in the town of Stewart. Let K be the number of days it will take for the plane to arrive in Phoenix for the first time from now. Determine the PMF of K.

$$p_K(k) = (k-1)(1-p_3)^2p_3^{k-2}, \quad k = 2, 3, ...$$

We note from the Markov chain that to travel from Stewart to Phoenix, transitions from Stewart to Colbert and from Colbert to Phoenix must occur. Both of these transition probabilities are $1 - p_3$. The other $k - 2$ transition probabilities are p_3. It is known that the last transition will be from Colbert to Phoenix, but the order of the other $k - 1$ transitions isn't set. Thus there are $\binom{k-1}{1} = k - 1$ ways in which we can pick when in the k-length sequence the transition from Stewart to Colbert occurs. Then it follows that

$$p_K(k) = (k-1)(1-p_3)^2p_3^{k-2}, \quad k = 2, 3, ...$$

Alternatively, K can be thought of as a sum of two independent geometric random variables (following the reasoning of the previous question) with probability of success $1 - p_3$. Thus, K is a Pascal random variable of order 2, and

$$p_K(k) = \binom{k-1}{2-1}(1-p_3)^2p_3^{k-2} = (k-1)(1-p_3)^2p_3^{k-2}, \quad k = 2, 3, ...$$

d. *(6 points)* The airline has been running this schedule for a long time. If the plane is in the air between Phoenix and Chicago, what is the probability that it is headed towards Chicago?

$$\mathbf{P}(\text{Plane is headed to Chicago}) = \frac{1}{2}$$

Since the airline has been running for a long time, we are in the steady state. Similarly to the argument for the balance equations in the birth/death processes, in the steady state, the frequency of travel from Phoenix to Chicago must be equal to the frequency of travel from Chicago to Phoenix. Thus the probability that the plane is heading towards Chicago is exactly $\frac{1}{2}$.

e. *(7 points)* In this question, assume that global warming suddenly makes the weather so bad that planes can land in small airports, but can no longer take off from small airports. The plane starts either in Chicago or in Phoenix, with equal probability. Find the probability that the plane started in Phoenix given it is now stuck in Stewart.

$\mathbf{P}(\text{started in Phoenix}|\text{stuck in Stewart}) = \frac{1}{1+p_2}$

Using Bayes' Rule,

$\mathbf{P}(\text{started in Phoenix}|\text{stuck in Stewart}) = \frac{\mathbf{P}(\text{stuck in Stewart}|\text{started in Phoenix})\mathbf{P}(\text{started in Phoenix})}{\mathbf{P}(\text{stuck in Stewart})}$

Let's now define the following absorption probabilities:

$$a_p = \mathbf{P}(\text{stuck in Stewart}|\text{started in Phoenix})$$
$$a_c = \mathbf{P}(\text{stuck in Stewart}|\text{started in Chicago})$$

From the Law of Total Probability,

$\mathbf{P}(\text{stuck in Stewart}) = \mathbf{P}(\text{started in Phoenix})a_p + \mathbf{P}(\text{started in Chicago})a_c = \frac{1}{2}a_p + \frac{1}{2}a_c$

and

$\mathbf{P}(\text{started in Phoenix}|\text{stuck in Stewart}) = \frac{\frac{1}{2}a_p}{\frac{1}{2}a_p+\frac{1}{2}a_c} = \frac{a_p}{a_p+a_c}$

It is now only necessary to find one relationship between a_c and a_p. From the Markov chain we notice that $a_c = p_2 a_p$, and

$\mathbf{P}(\text{started in Phoenix}|\text{stuck in Stewart}) = \frac{a_p}{a_p+a_c} = \frac{a_p}{a_p+p_2a_p} = \frac{1}{1+p_2}$

f. *(7 points)* In this question, again assume that global warming suddenly makes the weather so bad that planes can land in small airports, but can no longer take off from small airports. Also assume that the plane started in Chicago. Let M be the number of flights the plane made before it got stuck in a small town. Find the expectation of M.

$\mathbf{E}[M] = \frac{1+p_2}{1-p_1p_2}$

Let μ_c be the expected number of flights before the plane gets stuck in a small town, starting in Chicago. Also, let μ_p be the expected number of flights before the plane gets stuck in a small town, starting in Phoenix. Then by the Law of Total Expectation:

$\mu_c = 1 + \mu_p p_2 + 0 \cdot (1-p_2) = 1 + \mu_p p_2$

$\mu_p = 1 + \mu_c p_1 + 0 \cdot (1-p_1) = 1 + \mu_c p_1$

Solving this system of linear equations, we get:

$\mu_c = \frac{1+p_2}{1-p_1p_2}$, $\mu_p = \frac{1+p_1}{1-p_1p_2}$

Finally,
$\mathbf{E}[M] = \mu_c = \frac{1+p_2}{1-p_1p_2}$

g. *(7 points)* The airline has been running this schedule for a long time. It is known that on any given day, the plane is equally likely to be found in any of the six airports. Find the value of p_1, where $0.4 \leq p_1 \leq 0.6$, that maximizes the airline efficiency, that is, maximizes the fraction of the days the plane makes trip on average. Compute also this maximum efficiency.

Your Answer: $p_1 = 0.4$, efficiency $= 11/15$.

From the statement of the problem, during steady state, the probability of being in each of the states is $\frac{1}{6}$. Then,

(i) $\mathbf{P}(\text{being in Chicago}) = \frac{1}{6} = \mathbf{P}(\text{being in Shelbyville})(1-p_3) + \mathbf{P}(\text{being in Phoenix})p_1 = \frac{1}{6}(1-p_3) + \frac{1}{6}p_1$

(ii) $\mathbf{P}(\text{being in Phoenix}) = \frac{1}{6} = \mathbf{P}(\text{being in Colbert})(1-p_3) + \mathbf{P}(\text{being in Chicago})p_2 = \frac{1}{6}(1-p_3) + \frac{1}{6}p_2$

(iii) $\mathbf{P}(\text{being in Stewart}) = \frac{1}{6} = \mathbf{P}(\text{being in Phoenix})(1-p_1) + \mathbf{P}(\text{being in Stewart})p_3 = \frac{1}{6}(1-p_1) + \frac{1}{6}p_3$

(iv) $\mathbf{P}(\text{being in Springfield}) = \frac{1}{6} = \mathbf{P}(\text{being in Chicago})(1-p_2) + \mathbf{P}(\text{being in Springfield})p_3 = \frac{1}{6}(1-p_2) + \frac{1}{6}p_3$

(v) $\mathbf{P}(\text{being in Colbert}) = \frac{1}{6} = \mathbf{P}(\text{being in Stewart})(1-p_3) + \mathbf{P}(\text{being in Colbert})p_3 = \frac{1}{6}(1-p_3) + \frac{1}{6}p_3$

(vi) $\mathbf{P}(\text{being in Shelbyville}) = \frac{1}{6} = \mathbf{P}(\text{being in Springfield})(1-p_3) + \mathbf{P}(\text{being in Shelbyville})p_3 = \frac{1}{6}(1-p_3) + \frac{1}{6}p_3$

Equations (i) and (iii) imply that $p_3 = p_1$, while equations (ii) and (iv) imply that $p_2 = p_3$.

Therefore, $p_1 = p_2 = p_3$.

Airline efficiency would be maximized for the minimum possible p_3, since the efficiency is $1 - (1/6)(p_3)(4)$, but $p_1 = p_2 = p_3$ has to hold, and therefore the least possible value p_3 can take is 0.4, and the corresponding efficiency is 1-(1/6)(2/5)(4) = 11/15.

h. *(7 points)* In this question, assume that the plane starts in Phoenix and that $p_1 = p_2 = p_3 = \frac{1}{2}$. Let C be the event that by the time the plane takes off from the Stewart airport for the 50th time, it will have spent more than 90 nights at the Stewart airport. Find an *excellent* approximation to the probability of C.

You might find one of the following useful: $\sqrt{10} \approx 3.1623$, $\sqrt{100} = 10$.

$\mathbf{P}(C) \approx 0.8289$

Here we present two solutions to this problem. Both solutions are valid but lead to different approximations. The numerical answers assume that the argument of the standard normal CDF is rounded to two decimal points before using the standard normal CDF table.

Solution A:

Every time the plane lands in Stewart we create a random variable X_i that is geometric with parameter $p = \frac{1}{2}$. Then each X_i represents the number of nights spent in Stewart until departure. We find that $\mathbf{E}[X_i] = \frac{1}{p} = 2$ and $var(X_i) = \frac{1-p}{p^2} = 2$. Then the sum $\sum_{i=1}^{50} X_i$ represents the total number of nights spent in Stewart after 50 departures. Since X_i's are i.i.d., CLT applies. Using the half-step correction, we obtain:

$$\begin{aligned}
\mathbf{P}(C) &= \mathbf{P}\left(\sum_{i=1}^{50} X_i > 90\right) = \mathbf{P}\left(\frac{\sum_{i=1}^{50} X_i - 50(\mathbf{E}[X_i])}{\sqrt{50}\sigma_{X_i}} > \frac{90.5 - 50(\mathbf{E}[X_i])}{\sqrt{50}\sigma_{X_i}}\right) \\
&\approx 1 - \Phi\left(\frac{90.5 - 50(\mathbf{E}[X_i])}{\sqrt{50}\sigma_{X_i}}\right) = 1 - \Phi\left(\frac{90.5 - 100}{\sqrt{50}\sqrt{2}}\right) = 1 - \Phi(-0.95) = \Phi(0.95) \approx \boxed{0.8289}
\end{aligned}$$

Solution B:

Every night the plane spends in Stewart creates a random variable Y_i, that is Bernoulli with parameter $p = \frac{1}{2}$. We call a success the event that the plane flies out of Stewart the next day. We find that $\mathbf{E}[Y_i] = p = \frac{1}{2}$ and $var(Y_i) = p(1-p) = \frac{1}{4}$. Then the sum $\sum_{i=1}^{90} Y_i$ represents the total number of departures after spending 90 nights in Stewart. The plane spent more than 90 nights by the time the 50th departure takes place if and only if the total number of departures in the first 90 nights is less than 50. Since Y_i's are i.i.d., CLT applies. Using the half-step correction, we obtain:

$$\begin{aligned}
\mathbf{P}(C) &= \mathbf{P}\left(\sum_{i=1}^{90} Y_i < 50\right) = \mathbf{P}\left(\frac{\sum_{i=1}^{90} Y_i - 90(\mathbf{E}[Y_i])}{\sqrt{90}\sigma_{Y_i}} \leq \frac{49.5 - 90(\mathbf{E}[Y_i])}{\sqrt{90}\sigma_{Y_i}}\right) \\
&\approx \Phi\left(\frac{49.5 - 90(\mathbf{E}[Y_i])}{\sqrt{90}\sigma_{Y_i}}\right) = \Phi\left(\frac{49.5 - 45}{\sqrt{90}\frac{1}{2}}\right) \approx \Phi(0.9486) \approx \Phi(0.95) \approx \boxed{0.8289}
\end{aligned}$$

Note: We rounded 0.9486 to two decimal places to use the standard normal CDF table.