Problem 1 **(worth 15 pts)** This question asks you to compute the expectation of a geometric random variable X with probability $p = 0.25$. Since $E[X] = 1/p$, we have the expected # of attempts is $E[X] = 4$.

Problem 2 **((a) worth 15 pts, (b) worth 25 pts)** (a) This question asks you to compute the $P[Y = 8]$ where Y is a negative binomial random variable with probability $p = 0.80$ and number of successes $r = 6$. Since the formula for

$$P[Y = k] = \binom{k-1}{r-1} p^r (1-p)^{k-r},$$

plugging in r, k, and p as above, we get the probability that exactly 8 missiles destroy the target is $\binom{7}{5}(0.8)^6(0.2)^2 = 0.22020096$, or about 22%.

(b) The question could be done one of two ways. Firstly the probability that the gunboat (which has a stock of 10 missiles) destroys the target equals the sum

$$\sum_{k=0}^{10} P(exactly\ k\ missiles\ destroy) = \sum_{k=6}^{10} P(exactly\ k\ missiles\ destroy)$$

since we need at least 6 direct hits for destruction. Plugging in the formula from (a), we have

$$\begin{aligned}
\binom{5}{5}(0.8)^6(0.2)^0 &+ \binom{6}{5}(0.8)^6(0.2)^1 + \binom{7}{5}(0.8)^6(0.2)^2 \\
&+ \binom{8}{5}(0.8)^6(0.2)^3 + \binom{9}{5}(0.8)^6(0.2)^4 \\
&= 0.262144 + 0.3145728 + 0.22020096 \\
&+ 0.117440512 + 0.0528482304 \\
&= 0.9672065024 \approx 96.7\%
\end{aligned}$$

The second way to do this involves the Central Limit Theorem (This was recommended by the hint but it was prefectly valid to do using the negative binomial distribution).

As we saw in the Airplane Problem in the second computer lab, the Central Limit Theorem can be used as a tool to avoid the computation of a long sum, or other times where we do not necessarily have an effecient way to compute the probability exactly.

Following the hint, let Y_1 be the geometric random variable representing the amount of missiles needed to be fired the get the *first* direct hit. Similarly let Y_2 be the amount of missiles needed to get the next direct hit, etc. Thus the sum $Y_1 + Y_2 + \cdots + Y_6$ really represents the number of missiles need to get six direct hits. (Recall that a negative binomial random variable with parameters $p = 0.8$ and $r = 6$ is the sum of $r = 6$ independent geometric random variables with parameter $p = 0.8$.)

So using this reasoning, the probability that the gunboat destroys the target is precisely $P[Y_1 + Y_2 + \cdots + Y_6 \leq 10]$ where the $m = 10$ represents the maxinum number of missiles the boat can fire. Fortunately, the Central Limit Theorem gives a good way to approximate this probability, and in fact given that the expectaion of Y_i is $1/p = \mu = 1.25$ and the standard deviation of Y_i is $\frac{\sqrt{1-p}}{p} = \sigma = 0.5590169944$
Plugging these into the formula, we get

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{10-6(1.25)}{(0.559)\sqrt{6}}} e^{-t^2/2} dt$$

and after simplifying, we get this upper limit is 1.825741858.
To evaluate this integral, we look it up in Table A.1 on page 851 or we look at the table given at the bottom of the quiz for z_α. Since $z_{0.03} = 1.881$ and $z_{0.04} = 1.751$, we approximate that $z_\alpha = 1.826$ implies that $\alpha \approx 0.035$. Thus the area under the tail is about 0.035 and the desired integral is about $1 - 0.035 = 0.965$ agreeing (approximately) with the above answer.

Note: The length of this second explanation is a reflection on the fact that I am writing this as a solution set for you rather than the length of time it would take to solve and write-up the problem using this method. This method indeed only had three steps: calculating μ and σ using the formulas for a geometric random variables, plugging these into the upper limit to get 1.826 and looking up z_α on a table.

Problem 3 **(worth 15 pts)** We solve this problem by calculating z_{obs} by the formula

$$z_{obs} = \frac{\overline{\mu} - \mu_0}{\sigma/\sqrt{n}}$$

where $\overline{\mu}$ is the sample mean, n is the number of trials, μ_0 is the mean under the null hypothesis H_0, and σ is the standard deviation.
In this paricular example, H_0 is the hypothesis that $\mu = 450$, the sample mean is $\overline{\mu} = 461$, $n = 500$ California seniors, and $\sigma = 100$. Plugging that all in, we get $z_{obs} = 2.459674775$. (**This is the z-statistic**. Since we were testing this hypothesis versus $H_1 : \mu > 450$ at the 0.05 level of significance, we compare the z-statistic with $z_{0.05} = 1.645$ (from Table A.1 or given below.)
Since $z_{obs} > z_{0.05}$, means that the data of 461 is so high that we **reject** the null hypothesis of $\mu = 450$ (at the 0.05 level of significance).

Finally, the P**-value** for this problem is the probability that a second sample (obs′) would get data as skewed as this sample data was, i.e.

$$P[Z_{obs'} \geq z_{obs}] = P[Z_{obs'} \geq 2.459674775] = \frac{1}{\sqrt{2\pi}} \int_{2.459674775}^{\infty} e^{t^2/2} dt.$$

Again, you can't compute this integral by hand, so you should look up in table on page 851 to get P-value is approx. 0.007. Since this probability is so low, and more precisely, it is **less** than our threshold of 0.05, we reject. (Thus illustrating the two ways to conclude the same thing about this null hypothesis.)

Problem 4 **(worth 15 pts)** This is another hypothesis testing question, so the first thing we do is calculate z_{obs} using $\overline{\mu} = 11$, $\mu_0 = 12$, $n = 30$ and $\sigma = 3$. We get

$$z_{obs} = -1.825741858$$

where the negative sign means that the sample mean was less than the hypothesized mean (which comparing 11 and 12 should be no surprise.)
Again, we use the same logic as above since our alternative hypothesis H_1 is $\mu < 12$ (still one-sided – even though now picture is reflected).
Thus we compare the relative magnitudes (absolute values) and find that

$$1.825741858 > z_{0.05} = 1.645 \quad \text{so we again reject.}$$

This case the P-value is

$$\begin{aligned} P[Z_{obs'} \leq z_{obs}] = P[Z_{obs'} \leq -1.825741858] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-1.825741858} e^{t^2/2} dt \\ &= \frac{1}{\sqrt{2\pi}} \int_{1.825741858}^{\infty} e^{t^2/2} dt. \end{aligned}$$

where this last equality is by the symmetry of the bell-curve about the y-axis. (Notice the minus signs disappeared when we reflected.) Thus the P-value is approx. 0.034 (less than 0.05 so we reject.)

Problem 5 **(worth 15 pts)** We calculate z_{obs} using $\overline{\mu} = 0.4365$, $\mu_0 = 0.5$, $n = 100$ and $\sigma = \sqrt{\frac{1}{12}}$ and get

$$z_{obs} = -2.199704526.$$

Deciding whether to accept or reject takes a little more care this time however. Since are testing versus the alternate hypothesis $H_1 \mu \neq 1/2$, this is the double-sided case. So since we want area at the tails to add up to 0.02 (what we mean by level of significance) but we have both tails contributing (think of the outside of a confidence interval) this time, we really only compare with $z_{\alpha/2} = z_{0.01} = 2.326$ (wherer the last equality by chart we gave you. Tnus we check absolute values (like in problem 4) and get

$$2.199704526 = |z_{obs}| < z_{0.01} = 2.326 \quad \text{thus we } \textbf{do not reject}.$$

Note if you use $z_{0.02} = 2.054$ instead you would indeed conclude the wrong result from this test.
Finally, the P-value is

$$\begin{aligned} P[|Z_{obs'}| \geq |z_{obs}|] &= P[|Z_{obs'}| \geq 2.199704526] = 2 \cdot \frac{1}{\sqrt{2\pi}} \int_{2.199704526}^{\infty} e^{t^2/2} dt \\ &= 2 \cdot (0.01391393274) = 0.02782786548. \end{aligned}$$

Bonus **(worth 10 pts)** A uniform random variable on the interval $[0, 1]$ is a continous random variable with pdf given as

$$f(x) = \begin{cases} 1 & \text{if } 0 \leq \text{x} \leq 1 \\ 0 & \text{outside the unit interval.} \end{cases}$$

i.e. it's constant everywhere and has the property that $\int_{-\infty}^{\infty} f(x)dx = \int_0^1 f(x)dx = 1$. Thus you already have all the tools you need to calculate the mean and variance of this particular continuous random variable.

Namely,

$$\begin{aligned} \mu = E[X] &= \int_0^1 x \cdot f(x)dx = \int_0^1 xdx = 1/2, \quad \text{and} \\ E[X^2] &= \int_0^1 x^2 \cdot f(x)dx = \int_0^1 x^2dx = 1/3. \end{aligned}$$

Thus the Variance

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

resulting in the standard deviation of $\sqrt{Var[X]} = \sqrt{\frac{1}{12}}$.