1a: You must remove the bug before the $10^{th}$ trial. The (geometric) probability of this event is $\sum_{i=1}^{9} (1 - P)^{i-1}P$, where $P = 1/3$. This simplifies to $1 - (1 - P)^9 = .9739877$.

1b: This is the negative binomial probability that $W_3$, the wait for the third success, is less than 10. The expression is 
\[ (1/3)^3 \sum_{j=3}^{9} (2/3)^{j-3} \binom{j-1}{2}. \]

1c: Let $A$ be the event you remove the first bug on the second try. $P(A) = (2/3)(1/3) = 2/9$. Let $B$ be the event you remove the last (third) bug before, or ON trial 9. These events are independent because the trials are, so we seek $P_A(B) = P(B)$, and the latter is $P(W_2 = 2) + \cdots + P(W_2 = 7)$, i.e. $(1/3)^2 \sum_{n=2}^{7} (2/3)^{n-2} \binom{n-1}{1}$.

1d: The expected wait for two successes is $2P = 6$.

1e: The variance of the wait for $k$ successes is $k(1 - P)/P^2$. With our $k$ and $P$ this is 18.

2a: We did most of this question in class. $E(X) = E(Y) = 1$ and $E(Z) = 2$. $V(X) = V(Y) = 1/2$ and $V(Z) = 4$. $V(X + Y) = 0$ because $X + Y = 2$, a constant, and $V(X + Z) = 9/2$, which equals $V(X) + V(Z)$ so we know that Cov$(X, Z) = 0$; although these random variables are uncorrelated, we do NOT know so far whether or not they are independent.

2b: $\phi_X(s) = \phi_Y(s) = 1/4 + s/2 + s^2/4$ and $\phi_Z(s) = 1/2 + s^2/4$. $\phi_X(s) = 1/2 + s/2$ so $\phi_X'(1) = E(X) = 1$; same for $Y$. $\phi_Z'(1) = 2$.

2c: They are NOT independent because $\phi_X(s)\phi_Z(s)$ is $1/8 + s/4 + s^2/8 + s^4/8 + s^5/4 + s^6/8$ while $\phi_X+Z(s) = s/2 + s^2/4 + s^3/4$. The generating function of the SUM of independent, non-negative integer-valued random variables must be the product of their generating functions.

2c and 3: maybe later... for now they are too tedious.

**Question 4a:** We have $n = 1000$ Bernoulli trials where success is HEAD. If the coin were fair, we would expect $nP = 500$ successes, so from the given data, we have $t = 200$ too few. Tchebycheff’s inequality shows this event has probability at most $V(S_{1000})/t^2 = \frac{1000(1/2)^2(1/2)}{200^2} = 1/160$ (V($S_n$) = $nP(1 - P)$). We are $159/160 = 99.375\%$ confident the coin is biased against HEADS (there were too few).

**Question 4b:** We have $n$ Bernoulli trials where Success on a trial is “TAIL”. We are told that the number of successes is $S_n = n/3$. Recall that $E(S_n) = nP$ and $V(S_n) = nP(1 - P)$, so if the coin were fair, we expect $E(S_n) = n/2$ Tails and $V(S_n) = n/4$.

Tchebycheff’s inequality gives
\[ P(|S_n - n/2| \geq t) \leq \frac{V(S_n)}{t^2} = \frac{n/4}{t^2}; \]

We are told that $S_n = n/3$, which means that we observed a deviation of $t = n/6$ from the expected number of successes. According to the above, a deviation this large (or larger) would
occur with probability at most \(n/(4t^2) = 9/n\). The smaller this quantity, the more we believe the coin is biased against Tails.

In fact 1 minus the value of the right-hand side of the above inequality (in our case it is \(1 - 9/n\)) is the confidence for our conclusion, and we want this to be at least .95 (or 95 percent). Solving \(1 - 9/n > .95\) reveals that once \(n > 180\), we are 95 percent confident that we saw too few Tails to believe the coin is fair.

- **Question 4c:** The logic here is similar to 4b, above. \(S_{6000}\) is the number of threes that showed in 6000 tosses of a die. If the die were fair \(E(S_{6000}) = 1000\) and \(V(S_{6000}) = 5000/6\). Tchebycheff’s inequality gives

\[
P(|S_{6000} - 1000| \geq t) \leq \frac{5000}{6t^2}
\]

and to be 99 percent confident the deviation is large, we want the right hand side to be less than .01. This implies \(t^2 > 5000/06\), or that the deviation from the expectation is \(t > 288.675\ldots\). The required answer is \(1000 + 289\), where we added because the coin is biased in favor of THREE, so we need to observe MORE than expected.

- **Question 6:** The number of \(n\)-node rooted binary trees is \(b_n = \frac{1}{n+1}\binom{2n}{n}\). For \(n = 7\) this is 429.

25 have three nodes in the left subtree because if there are three nodes on the left subtree, there are \(b_3 = 5\) different left subtrees. For each there must also be three nodes in the right subtree (since there are seven nodes overall), so again 5 different right subtrees. By the cartesian product principle, 25 seven-node rooted trees have three nodes on the left.

- **Question 7 (harder):** Let \(A\) be the event that the random tree has three nodes in the left subtree, and \(B\) the event it has height 3. \(P_A(B) = P(A \cap B)/P(A) = |A \cap B|/|A|\). We know from problem 1 that \(|A| = 25\). \(B\) occurs unless both left and right (3-node) subtrees are balanced, so \(|A \cap B| = 24\).