Problem 1.

(a) There are multiple related ways of proving this claim: use the known results on gambler's ruin, or write a recurrence relation for the probability of reaching \( b \) before \(-a\) starting at \( k \in \{a, a+1, \ldots, b\} \). It is also possible to argue that the sequence \( (S_n)_n \) forms a martingale w.r.t. the natural filtration. The conditions for finite expectation \( E[S_{\infty}] \) and \( E[S_n \mid S_{n-1}] \) are checked immediately.

Then observe that the RV \( \tau = \min\{ t \in \mathbb{N} : S_t = -a, S_{t+1} \neq b, S_{t+2} \neq b, \ldots, S_{t+k} \neq b \} \) is a stopping time for the sequence \( (S_n)_n \).

The optional stopping theorem for the stopped martingale \( S_\tau \) implies that \( E[S_\tau] = E[S_0] = 0 \). At the same time,

\[
E[S_\tau] = -a(1-p) + b p, \quad \text{where } p \text{ is the probability of reaching } b \text{ before reaching } -a; \text{ thus } p = \frac{a}{a+b} \quad \text{(we used this argument in class; Lect. 12)}
\]

(b) \[ E \left[ (S_{n+1} + Z_{n+1})^2 - (n+1) \mid S_n \right] = E \left[ S_n^2 + 2S_nZ_{n+1} + Z_{n+1}^2 - (n+1) \mid S_n \right] \]

\[ = S_n^2 + 2S_n E[Z_{n+1}] + 1 - n - 1 = S_n^2 - n. \]

Since \( E[S_n^2] = E[S_T^2 - T] = 0 \) (again using the optional stopping theorem), we obtain \( ET = E[S_T^2] = a^2 \).

(c) This is a direct calculation:

\[
E \left[ (S_n + Z_{n+1})^4 - 6n (S_n + Z_{n+1})^2 + b(n+1)^2 + c(n+1) \mid Y_n \right] \leq Y_n
\]

This yields \( b = 3, c = 2 \).

Next,

\[
0 = E \left[ S_T^4 - 6TS_T^2 + 3T^2 + 2T^2 \right] = E \left[ a^4 - 6Ta^2 + 3T^2 + 2T \right]
\]

\[ = 3ET^2 - 5a^4 + 2a \]

\[
\Rightarrow ET^2 = \frac{a}{3}(5a^3 - 2)
\]
Problem 2.

1. The generating function of the RV $Z$ that is 0 with prob. $1-p$ and 2 w. prob. $p$ equals
   \[ g_Z(s) = 1 - p + p s^2 \]
   Compute $g'_Z(s) = 2ps$; $g''_Z(s) = 2p$ to conclude that $g(s)$ is a convex increasing function.
   The extinction probability of the GW process is given by the fixed point of $g_Z(s)$.
   If $p < \frac{1}{2}$, then the only fixed point is $s = 1$, and thus
   \[ \mathbb{P}_1 = P(E) = 1 \]
   If $p > \frac{1}{2}$ then from $\mathbb{P} = g_Z(\mathbb{P})$ we find
   \[ \mathbb{P} = \frac{1-p}{p} \]
   as required.

2. This follows more or less by definition. The idea is that $Y_i = 1$ corresponds to adding 2 children of node $i$ while $Y_i = -1$ corresponds to adding none.

   Let $T$ be a GW tree. With the above, the number of leaves in level $i$ is
   \[ X_i = 1 + \sum_{j=1}^{i} Y_i \]
   where 1 is added to account for the root (i.e. $X_0 = 1$).
   Thus, $X_{i-1} = S_i$, and if $S_i = -1$, the process becomes extinct, i.e., stops, and once it stops, it is stopped forever, implying that $S_j > -1$ for all $j < i$.
   Thus
   \[ P(W = k) = P(S_k = -1, \text{ and } S_i > 0 \text{ for all } 1 < i < k) \]
   By a change of variable, let us start the walk at -1 and assume that the first step is to the right, which happens with prob. $p$. We obtain
   \[ P(W = k) = P(S_0 = S_k = 0, \text{ and } S_i > 0 \text{ for all } 1 < i < k) \]
Problem 3

(a) We need to show that for any $p > \frac{1}{2}$ there is a $\lambda > 0$ that satisfies the equality

$$E[Y_n | X_{n-1}] = Y_{n-1}$$

with $Y_n = e^{-\lambda X_n}$, $n \geq 1$. [It is clear that $E[Y_n] < \infty$ for all $n$.]

Since $E e^{-\lambda X_n}$ is an m.g.f. of the RV $X_n \sim \text{Binom}(2X_{n-1}, p)$, we obtain

$$E[Y_n | X_{n-1}] = (p e^{-\lambda} + 1 - p)^{2X_{n-1}}.$$  

For this to be equal to $Y_{n-1} = e^{-\lambda X_{n-1}}$, it is necessary and sufficient that

$$(p e^{-\lambda} + 1 - p)^2 = e^{-\lambda}$$

This is a quadratic equation for $e^{-\lambda}$ which yields $e^{-\lambda} = \left(\frac{1-p}{p}\right)^2$ or

$$\lambda = 2 \log \frac{p}{1-p}.$$  

This is positive if $p > \frac{1}{2}$.

(b) $(Y_n)_n$ is a bounded martingale, so it converges a.s. and in $L^1$.

Since $0 \leq Y_n \leq 1$ for all $n$, also $0 \leq Y_{\infty} \leq 1$. We also have $E Y_{\infty} = e^{-\lambda}$, and thus, $P(Y_{\infty} = 1) \leq e^{-\lambda} < 1$ for any $\lambda > 0$, or $p > \frac{1}{2}$.

Further, $\{X_n = 0\} = \{Y_n = 1\}$. If the open cluster extends to infinity with positive probability, $P(X_n = 0) < 1$ for all $n$. Since $P(Y_{\infty} = 1) < 1$, also $P(X_{\infty} = 0) < 1$, and thus there is an $\varepsilon$ s.t. $P(X_n = 0) < 1 - \varepsilon$ for all $n$.

Since $\lim_{n \to \infty} P(Y_n = 1) < 1$, there is an $\varepsilon > 0$ s.t. $P(Y_n = 1) < 1 - \varepsilon$ for all $n$.

This implies that $P(X_n = 0) < 1 - \varepsilon$ for all $n$, and this in turn implies that with positive probability, $X_n > 0$ for all $n$, i.e., there exists an infinite cluster.

Thus, for any $p > \frac{1}{2}$, percolation occurs, i.e., $p_c \leq \frac{1}{2}$.  


Problem 4

We need some estimate of \( \binom{n}{m} \). You can use
\[
\binom{n}{m} \approx \frac{n^m}{2n^{m-n} \left( \frac{m}{n} \right)^m \left( 1 - \frac{m}{n} \right)^{n-m}}
\]
(Wikipedia).

We will use a slightly more precise one
\[
\binom{n}{m} \geq \frac{1}{\sqrt{8m(1 - \frac{m}{n})}} \left( \frac{m}{n} \right)^m \left( 1 - \frac{m}{n} \right)^{n-m} \quad \text{(this is responsible for } \sqrt{8} \text{)}
\]

Since \( P_x(x) = \binom{n}{m}^{-1} \) and \( Q_x(x) = \left( \frac{m}{n} \right)^n \left( 1 - \frac{m}{n} \right)^n \), we obtain for any \( x \)
\[
\frac{P_x(x)}{Q_x(x)} = \binom{n}{m}^{-1} \left( \frac{m}{n} \right)^n \left( 1 - \frac{m}{n} \right)^n \leq \sqrt{8} \sqrt{\frac{m(n-m)}{n}}.
\]

We note that a stronger version of this inequality, due to Hoeffding, enables one to replace the constant on RHS by 1 (see [FKL+], Lemma 21.1).

Thus
\[
E_p e^{\lambda f(x)} = \sum_{x} P_x(x) e^{\lambda f(x)} \leq \sqrt{8} \sqrt{\frac{m(n-m)}{n}} E_q e^{\lambda f(x)}
\]

Now all we need to prove is
\[
(A) \quad E_q e^{\lambda f(x)} \leq e^{2n\lambda^2 + \lambda E_q(f)}
\]

The bounded differences inequality implies that
\[
P(\left| f(x) - EF(x) \right| \geq t) \leq 2e^{-2t^2/n}, \quad \text{for all real } t,
\]

i.e., \( f(x) \) is \((2,n)\) subgaussian.

Assume for the moment that \( EF(X) = 0 \). We have
\[
E_q e^{\lambda f(x)} = \int_0^\infty P(e^{\lambda f(x)} > s) ds = \int_0^\infty P(f(x) > \frac{\lambda s}{x}) ds = \int_0^\infty e^{\lambda u} P(f(x) > \frac{u}{\lambda}) du
\]
\[
\leq 2 \int_0^\infty e^{-u^2/2n\lambda^2} du = 2e^{-n\lambda^2} \int_0^{\infty} e^{-u^2/2n\lambda^2} du = 2 \sqrt{2\pi n\lambda^2} e^{-n\lambda^2/2}
\]
\[
\leq e^{2n\lambda^2}
\]

For \( EF(X) \neq 0 \) we get the additional factor \( e^{\lambda E_q(f)} \), which establishes (A).
Problem 5

(a) Following the hint, we have
\[ 1 = (a + 1 - a)^n \leq \sum_{i=0}^{n} \left( \begin{array}{c} n \\ i \end{array} \right) a^i (1-a)^{n-i} \leq \sum_{i=0}^{n} \left( \begin{array}{c} n \\ i \end{array} \right) a^i (1-a)^{n-i} \left( \frac{a}{1-a} \right)^{i} \sum_{i=0}^{n} \left( \begin{array}{c} n \\ i \end{array} \right) \]
where for the last step we assumed that \( a < \frac{1}{2} \).

This shows that not only \( \left( \begin{array}{c} n \\ i \end{array} \right) < e^{n \log(a)} \) but even \( \sum_{i=0}^{n} \left( \begin{array}{c} n \\ i \end{array} \right) < e^{n \log(a)} \), \( a < \frac{1}{2} \).

Since \( \left( \begin{array}{c} n \\ i \end{array} \right) = \left( \begin{array}{c} n \\ n-i \end{array} \right) \), we can similarly prove the claim for \( a > \frac{1}{2} \); for \( a = \frac{1}{2} \) the inequality is trivial: \( \left( \begin{array}{c} n \\ i \end{array} \right) \leq 2^n \).

(b) The probability in question equals \( \sum_{i=\alpha n}^{n} \left( \begin{array}{c} n \\ i \end{array} \right) p^i (1-p)^{n-i} \). Since \( \alpha > p \), we calculate
\[ z := \frac{\alpha}{1-\alpha} \frac{1-p}{p} > 1. \]
Then
\[ \sum_{i=\alpha n}^{n} \left( \begin{array}{c} n \\ i \end{array} \right) p^i (1-p)^{n-i} \leq \sum_{i=\alpha n}^{n} \left( \begin{array}{c} n \\ i \end{array} \right) \frac{p}{1-\alpha} \left( \frac{1-p}{p} \right)^{n-i} \]
\[ = z^{-\alpha n} (pz + 1-p)^n \leq z^{-\alpha n} \left( \frac{p}{1-\alpha} \right)^{\alpha n} \left( \frac{1-p}{p} \right)^{-\alpha n} \]
\[ = e^{-nD(a \| p)} \]

Next, use the Cramér–Chernoff bound. Let \( X \sim \text{Binom}(n,p) \)
\[ \psi_X(\lambda) = n \log(e^\lambda p + 1-p) \]
Solving the equation \[ \psi_X'(\lambda) = \frac{np e^\lambda}{e^\lambda p + 1-p} = t \] (which gives \( \lambda \) that maximizes \( \lambda t - \psi_X(\lambda) \))
we find \( \lambda^* = \log \left( \frac{(1-p)t}{p(n-t)} \right) \)
Then \( P(X > \alpha n) = \exp \left( - \psi_X(\alpha n) \right) = \exp \left( n \left( \alpha \log \frac{1-p}{p} + \log \frac{1-p}{1-\alpha} \right) \right) \)
\[ = \exp \left( - n \log \frac{\alpha}{p} - (1-\alpha) n \log \frac{1-\alpha}{1-p} \right) = e^{-nD(\alpha \| p)} \]
\[ = e^{-2 \log \frac{\alpha}{p}} \]
If \( \alpha = \frac{b}{n} \) and \( p = \frac{c}{n} \), then Hoeffding's inequality gives \( e^{-\frac{2(b-c)^2}{n}} \)
which is a weak (for \( n \to \infty \), a trivial) estimate.
Problem 5, part (c)

Let us compute $EY$, where $Y = |\{v \in V : \deg(v) > b\}|$ is the random number of vertices of degree $> b$

$$EY = \sum_{v \in V} E[1_{\{\deg(v) > b\}}] = n \sum_{i=0}^{n} \binom{n-1}{i} p^i (1-p)^{n-i} \cdot 1_{\{\deg(v) > b\}}$$

$$\leq n \exp \left( -(n-1) \left( \frac{b}{n-1} \log \frac{b/n-1}{p} + (1- \frac{b}{n-1}) \log \frac{1-b/n-1}{1-p} \right) \right)$$

$$= n \exp \left( -b \log \frac{b}{c} \cdot \frac{n}{n-1} - (n-1-b) \log \frac{n-1-b}{(1-c/n)(n-1)} \right)$$

$$\leq n \exp \left( -b \log \frac{b}{c} - (n-1-b) \log \frac{n-1-b}{(n-1)(n-1)} \right)$$

$$\leq n \exp \left( -b \log \frac{b}{c} - (n-1) \log \frac{n-1-b}{n-c} + b \log \frac{n-1-b}{n-c} \right)$$

$$= n \exp \left( -b \log \frac{b}{c} - (n-1) \log \frac{n-1-b}{n-c} + b \log \frac{n-1-b}{n-c} \right)$$

While $\log \frac{n-1-b}{n-c} = \log \left( \frac{n-1-b}{n-c} - \frac{1}{n-c} \right) > \log (1 + \frac{b-c-1}{n-c}) + \mathcal{O}(n^{-1}) = \frac{b-c-1}{n-c} + \mathcal{O}(n^{-1})$

Now continue as follows:

$$\leq n \exp \left( -b \log \frac{b}{c} - (n-1) \frac{b-c-1}{n-c} + \mathcal{O}(n^{-1}) \right) \leq n \exp \left( -b \log \frac{b}{c} - c + b + \mathcal{O}(n^{-1}) \right)$$

Using the first moment method,

$$P(\Delta(G) > b) = P(Y > 0) \leq EY < \infty$$

Now let us bound above $EY^2$, envisioning the use of the second moment method.

$$EY^2 = \sum_{v,v'} E[1_{\{\deg(v) > b\}} 1_{\{\deg(v') > b\}}] = EY + \sum_{v \neq v'} E[1_{\{\deg(v) > b\}} 1_{\{\deg(v') > b\}}]$$
Bounding the 2nd term on the right-hand side of the above (last) expression, let
\[ y := \exp(-b \log \frac{b}{c} + b + O(n^{-1})) \]

The probability under the sum \( \sum_{u \neq u'} \) depends on the connection status of \( u \) and \( u' \); for each summand we have
\[ P(\text{deg}(u) > b, \text{deg}(u') > b) = p^2 P(\text{Binom}(n-2, p) > b-1)^2 + (1-p)^2 P(\text{Binom}(n-1, p) > b)^2 \leq y^2. \]

Then \( EY^2 \leq nY + n^2Y^2 = nY(nY+1) \)

and
\[ P(\Delta(G) > b) = P(Y > 0) \geq \frac{(EY)^2}{EY^2} = \frac{n^2Y^2}{nY + n^2Y^2} = \frac{1}{1 + \frac{1}{nY}} \]

which is the claimed expression.

(d) This part is proved in the book [FK16], Thm 3.4 which shows a stronger result:
\[ \lim_{n \to \infty} P(\Delta(G) = \log n / \log \log n) = 1. \quad (F) \]

Many students copied their proof, but we didn't suffer for nothing in parts (b)-(c). Indeed, to prove (F) above, we take the estimates established in those parts and observe that they can be used to show that
\[ P(\Delta(G) > b) \to 0 \quad \text{if} \quad b = \frac{\log n}{\log \log n} + o_1(n) \]
\[ P(\Delta(G) > b) \to 1 \quad \text{if} \quad b = \frac{\log n}{\log \log n} - o_2(n) \]
Problem 6

Part 1: The operator $D_i$ is a linear operator, $D_i(af + g) = aD_i(f) + D_i(g)$. We write $f$ using the Fourier expansion and apply $D_i$ to obtain

$$D_i(f) = D_i \left( \sum_{S \subseteq [n]} \hat{f}(S) \chi_S \right) = \sum_{S \subseteq [n]} \hat{f}(S) D_i(\chi_S).$$

The result follows by noticing that if $i \notin S$ then $D_i(\chi_S) = 0$ and if $i \in S$ then $D_i(\chi_S) = \chi_{S \setminus \{i\}}$.

By the definition of $I_i(f) = \Pr(f(X) \neq f(X^{(i)}))$, if $1_f$ is the indicator for $f(X) \neq f(X^{(i)})$ then $I_i(f) = \mathbb{E}[1_f]$. But if $f(X) = f(X^{(i)})$ then $D_i(f) = 0$ and if $f(X) \neq f(X^{(i)})$ then $D_i(f) \in \{1, -1\}$. Therefore, $(D_i(f))^2$ is exactly the indicator $1_f$, which implies $I_i(f) = \mathbb{E}[(D_i(f))^2] = \|D_i f\|^2_2$.

Part 2: Again we use the definition and obtain

$$\mathbb{E}[\|\nabla f\|^2_2] = \mathbb{E}[\langle \nabla f, \nabla f \rangle] = \mathbb{E}\left[ \sum_{i=1}^n (D_i f)^2 \right] = \sum_{i=1}^n I_i(f) = I(f).$$

Part 3: We saw in class that

$$\mathbb{E}[\|\nabla f\|^2_2] = I(f) = \sum_{S} |S|^2 \hat{f}(S)^2$$

and that

$$\text{Var}(f) = \sum_{S \neq \emptyset} \hat{f}(S)^2$$

and the result follows.

Problem 7

Part 1: First notice that $T_\rho f$ is linear (since the expectation is linear) and recall from class that $T_\rho(\chi_S) = \rho^{|S|} \chi_S$. Now we write

$$T_\rho f = T_\rho \left( \sum_{S \subseteq [n]} \hat{f}(S) \chi_S \right) = \sum_{S \subseteq [n]} \hat{f}(S) (T_\rho \chi_S) = \sum_{S \subseteq [n]} \hat{f}(S) \rho^{|S|} \chi_S.$$ 

For functions $f, g$, using the Fourier expansion we obtain

$$S_\rho(f, g) = \langle f, T_\rho g \rangle = \sum_{S \subseteq [n]} \sum_{B \subseteq \rho^{|S|}} \hat{f}(S) \rho^{|B|} \hat{g}(B) \langle \chi_S, \chi_B \rangle = \sum_{S \subseteq [n]} \hat{f}(S) \rho^{|S|} \hat{g}(S).$$

Part 2: Note that

$$S_\rho(f) = \mathbb{E}[f \cdot T_\rho f] = \langle f, T_\rho f \rangle = \Pr(f = T_\rho f) - \Pr(f \neq T_\rho f) = 2 \Pr(f \neq T_\rho f) - 1.$$ 

Taking $\rho = 1 - 2\epsilon$ we obtain that $y_i = x_i$ with probability $\frac{1}{2} + \frac{1}{2} \epsilon = 1 - \epsilon$ and $y_i = -x_i$, with probability $\epsilon$. Therefore,

$$\frac{1}{2} - \frac{1}{2} S_{1-2\epsilon}(f) = \frac{1}{2} - \frac{1}{2} (2 \Pr(f \neq T_{1-2\epsilon} f) - 1) = 1 - \Pr(f \neq T_{1-2\epsilon} f) = \Pr(f = T_{1-2\epsilon} f) = NS_\epsilon(f).$$

Part 3: First, we show that $T_\rho \chi_S = \rho^{|S|} \chi_S$. Indeed,

$$T_\rho \chi_S = \mathbb{E}_{y \sim \chi_S \mid \chi_S(y]} = \mathbb{E}_{y \sim \chi_S \mid \prod_{i \in S} y_i} = \prod_{i \in S} \mathbb{E}_{y \sim \chi_S \mid y_i} = \prod_{i \in S} \rho x_i = \rho^{|S|} \chi_S.$$ 

Now we write

$$S_\rho(D_i f) = \langle D_i f, T_\rho D_i f \rangle = \langle \sum_{S \ni i} \hat{f}(S) \chi_S \setminus \{i\}, T_\rho \left( \sum_{B \ni i} \hat{f}(B) \chi_B \setminus \{i\} \right) \rangle = \langle \sum_{S \ni i} \hat{f}(S) \chi_S \setminus \{i\}, \sum_{B \ni i} \hat{f}(B) T_\rho \chi_B \setminus \{i\} \rangle = \langle \sum_{S \ni i} \hat{f}(S) \chi_S \setminus \{i\}, \sum_{B \ni i} \hat{f}(B) \rho^{|B| - 1} \chi_B \setminus \{i\} \rangle = \sum_{S \ni i} \sum_{B \ni i} \hat{f}(S) \hat{f}(B) \rho^{|B| - 1} \chi_S \setminus \{i\} \cdot \chi \setminus \{i\} = \sum_{S \ni i} \hat{f}(S) \rho^{|S| - 1}.$$
Problem 8

part 1: First, it is clear (telescopic sum) that \( V = \sum_{i=1}^{n} V_i \). From the tower property of the expected value we have that for \( j < i \),
\[
\mathbb{E}[\mathbb{E}[Z \mid X_1, \ldots, X_i] \mid X_1, \ldots, X_j] = \mathbb{E}[Z \mid X_1, \ldots, X_j].
\]
Therefore, for \( j < i \) we obtain
\[
\mathbb{E}[V_i \mid X_1, \ldots, X_j] = \mathbb{E}[\mathbb{E}[Z \mid X_1, \ldots, X_i] \mid X_1, \ldots, X_j] - \mathbb{E}[\mathbb{E}[Z \mid X_1, \ldots, X_{i-1}] \mid X_1, \ldots, X_j] = 0.
\]
This implies that for \( j < i \)
\[
\mathbb{E}[V_i V_j] = \mathbb{E}[\mathbb{E}[V_i \mid X_1, \ldots, X_j]] = 0.
\]

part 2:
\[
V_i^2 = (\mathbb{E}[Z \mid X_1, \ldots, X_i] - \mathbb{E}[Z \mid X_1, \ldots, X_{i-1}])^2
\]
\[
= (\mathbb{E}[\mathbb{E}[Z \mid X_1, \ldots, X_n] - \mathbb{E}[Z \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n] \mid X_1, \ldots, X_i])^2
\]
\[
\leq \mathbb{E}[(\mathbb{E}[Z \mid X_1, \ldots, X_n] - \mathbb{E}[Z \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n])^2 \mid X_1, \ldots, X_i]
\]
\[
= \mathbb{E}[(Z - \mathbb{E}[Z])^2 \mid X_1, \ldots, X_i].
\]

part 3: We now calculate the variance of \( Z \).
\[
\text{Var}(Z) = \mathbb{E} \left( \sum_{i=1}^{n} V_i \right)^2
\]
\[
= \mathbb{E} \left[ \sum_{i=1}^{n} V_i^2 \right] + 2 \mathbb{E} \left[ \sum_{j<i} V_i V_j \right]
\]
\[
= \mathbb{E} \left[ \sum_{i=1}^{n} V_i^2 \right].
\]
From Part 2 we obtain
\[
\mathbb{E}[V_i^2] \leq \mathbb{E}[(Z - \mathbb{E}[Z])^2 \mid X_1, \ldots, X_i] = \mathbb{E}[(Z - \mathbb{E}[Z])]^2.
\]
Thus, we obtain
\[
\text{Var}(Z) \leq \sum_{i=1}^{n} \mathbb{E}[(Z - \mathbb{E}[Z])^2].
\]