Problem 1. Consider percolation on a graph with vertex set $\mathbb{N}$ with m_n parallel edges between n and $n+1$.
(a) Determine a necessary and sufficient condition on the sequence (m_n) for which there is an infinite connected cluster.
(b) Use this to find a sequence so that $p_c = 1/2$ and at $1/2$ there is a.s. an infinite cluster.

Solution. Let A_n be the event that at least one edge between n and $n+1$ is open. Then an infinite cluster exists if and only if A_n occurs eventually (efor all $n > n_0$). By Borel-Cantelli, this is equivalent to $\sum \mathbb{P}(A_n^c) < \infty$. We have $\mathbb{P}_p(A_n^c) = (1-p)^{m_n}$, so the condition for percolation is $\sum(1-p)^{m_n} < \infty$.

Take m_n so that $2^{m_n} \asymp n \log^2 n$. (Here $\asymp$ means equal up to some absolute constants.) Then at $p = 1/2$ there is percolation since $\sum \frac{1}{n \log^2 n} < \infty$. This means $m_n \sim \log n / \log 2$. For any smaller $p < 1/2$ we have $(1-p)^{m_n} = (2^{m_n})^{-\log(1-p)/\log(2)}$ tends to 0 like a smaller power of n, so $\sum(1-p)^{m_n} = \infty$ and there is no percolation.

Note: the condition needed here is that $q_n = 1/2^{m_n}$ is summable, but for any $\varepsilon > 0$ we have $\sum q_n^{1-\varepsilon} = \infty$.

Problem 2. Let $X_1, X_2, \dots$ be i.i.d. random variables with standard normal distribution, and $M_n = \max(X_1, \dots, X_n)$.
(a) Find $\limsup \frac{X_n}{\sqrt{\log n}}$.
(b) Show that $M_n/\sqrt{\log n}$ converges a.s. to the same value. (Hint: You need some estimates on the tail of the normal distribution.)

Solution. Let $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ be the density of X_i.

Lemma 1. *For $t > 1$ we have that $\frac{1}{2t} f(t) \le \mathbb{P}(X > t) \le \frac{2}{t} f(t)$.*

Proof. For $x > t$, the density $f(x)$ is above the tangent line at t, and below $\frac{1}{\sqrt{2\pi}} e^{-xt/2}$. Integrating over $[t, \infty)$ gives the claim. □

- It follows from Borel-Cantelli that $\limsup X_n/\sqrt{\log n} = \sqrt{2}$ a.s.: We have $\sum \mathbb{P}(X_n > a\sqrt{\log n} < \infty$ for $a > \sqrt{2}$ and the sum is infinite for $a < \sqrt{2}$.
- To get te result for M_n, let $Y_i = X_i/\sqrt{\log i}$. Fix some k, and note that

$$\frac{M_n}{\sqrt{\log n}} \le \max\left(\frac{X_1}{\sqrt{\log n}}, \dots, \frac{X_k}{\sqrt{\log n}}, Y_{k+1}, \dots, Y_n\right).$$

Now, the first k terms tend to 0, and $\limsup Y_n = \sqrt{2}$, so $\limsup M_n \le \sqrt{2}$.
For a matching lower bound, use that for large n, $M_n/\sqrt{\log n} \ge (1-\varepsilon)\max\left(Y_{n/2}, \dots, Y_n\right)$. An easy application of Borel-Cantelli gives $\liminf M_n/\sqrt{\log n} \ge (\sqrt{2} - \varepsilon)$.

Problem 3. Show that for any random variable with finite mean, $t\mathbb{P}(X > t) \to 0$ as $t \to \infty$. Moreover, if $\mathbb{E}|X|^a < \infty$ then $t^a \mathbb{P}(X > t) \to 0$.

Solution. Let $Y_t = X 1_{X>t}$. Then $Y_t \to 0$, and $Y_t \le |X|^a$. By dominated convergence $\mathbb{E}Y_t \to 0$. However, $\mathbb{E}Y_t \ge t^a \mathbb{P}(X > t)$, giving the claim.

Problem 4. Let X_i be i.i.d. random variables, and $S_n = \sum_{i \le n} X_i$.
(a) Show that if $\mathbb{E}X_i = +\infty$ then $\frac{1}{n} S_n \to \infty$ a.s.
(b) Show that if $\mathbb{E}X_i$ is not defined (positive and negative parts both infinite) then a.s. $\limsup \frac{1}{n} S_n = \infty$ and $\liminf \frac{1}{n} S_n = -\infty$.

Solution.

- For any $M > 0$, let $Y_i = X_i \wedge M$. Then $\mathbb{E}Y$ is finite and tends to infinity as $M \to \infty$. Thus for any K we can find M so that $\mathbb{E}Y > K$. Let $T_n = \sum_{i \le n} Y_i$. By the LLN, $T_n/n \to \mathbb{E}Y$, and so $T_n/n > K$ eventually. The claim follows since $S_n \ge T_n$.
- This is wrong as written. The $\limsup S_n/n$ is a tail random variable: not affected by changing any finite number of X_is. By Kolmogorov, $\limsup S_n/n$ must be trivial (always take the same value).
  To show that the $\limsup S_n/n$ cannot be se cannot be any finite value,

Definition 1. A random variable has density f if it's distribution function is $F(t) = \int_{-\infty}^t f(x)dx$. (In that case, $f(x) = F'(x)$.) A random variable with a density is called continuous.

Problem 5. If X is a continuous random variable with distribution function F and density $f = F'$, show (from the definition of expectation in class) that $\mathbb{E}X = \int_{-\infty}^{\infty} xf(x)dx$.

Solution. For $X > 0$, take a simple random variable $Y_n = [nX]/n$, where $[x]$ is the integer part (that is, X rounded down to a multiple of $1/n$). We have by definition

$$\mathbb{E}Y_n = \sum_i \frac{i}{n}\mathbb{P}(i \le nX < i+1) = \int_0^\infty [nx]/nf(x)dx.$$

This is a Riemann sum for the integral, and as $n \to \infty$ this approaches $\int_0 xf(x)dx$.

Any other simple variable $Z \le X$ satisfies $Z \le Y + 1/n$, so the expectation of X cannot be larger.

For non-positive X, the density of X^+ is f on $\mathbb{R}_+$, and of X^- it is $f(-x)$, so we can write the integral as $\int_{-\infty}^0 + \int_0^\infty$.

Definition 2. A pair of random variables X, Y has join density function $f : \mathbb{R}^2 \to \mathbb{R}$ if $\mathbb{P}((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy$.

Problem 6. Suppose continuous random variables X_1, X_2 have densities f_1, f_2 respectively. Prove from the definition of independence that they are independent if and only if they have a joint density of the form $f(x,y) = f_1(x)f_2(y)$.

Solution. For intervals I_1, I_2, we have $\mathbb{P}(\forall i, X_i \in I_i) = \iint_A f(x,y)\,dx\,dy$. If f factors, then this is a product of $\mathbb{P}(X_i \in I_i)$, and so the events $\{X_i \in I_i\}$ are independent. Since this is true for intervals, the independence extends to arbitrary measurable sets A_i.

Conversely, if the variables are independent, let $I_i = [x_i, x_i + \varepsilon]$. Then from the definition of joint density $\mathbb{P}(\forall i, X_i \in I_i) \sim \varepsilon^2 f(x_1, x_2)$. On the other hand by independence this is close to $\varepsilon f_1(x_1) \cdot \varepsilon f_2(x_2)$. Equating the two gives the factorization of f.

Problem 7. Let X, Y are two independent random variables. For an angle θ, the rotation by theta is the pair of random variables

$$X' = X\cos\theta + Y\sin\theta, \qquad Y' = -X\sin\theta + Y\cos\theta.$$

Find a joint distribution of independent variables X, Y, with $\mathbb{P}(X = 0) < 1$ so that for any θ the rotation (X', Y') has the same (joint) distribution as (X, Y).
Bonus: Find (with proof) all such distributions.

Solution. i.i.d. standard Gaussian X, Y have density $f(x,y) = \frac{1}{2\pi}e^{-(x^2+y^2)/2}$. This is rotationally symmetric, so is invariant to rotations.

If the variables have finite variance σ^2, the following shows uniqueness: Let S_N with $N = 2^n$ be a sum of 2^n i.i.d. copies of X. We have that $(X + X')/\sqrt{2}$ has the same law as X, so S_2 has the same distribution as $\sqrt{2}X$. By induction, S_N has the same law as $\sqrt{N}X$, and therefore $S_N/\sqrt{N}$ has the same law as X. However, the CLT implies $S_N/\sqrt{N} \to N(0, \sigma^2)$, so X must be Gaussian.

Without assuming finite variance the proof is more delicate.

***Problem 8.** (a) For given $p \in (0, 1)$, construct a tree where $p_c = p$.

(b) Find a tree where $p_c = p$ and $\theta(p_c) > 0$.

(Hint: if a vertex in level n has d_n children, write recursions for the probability that 0 is connected to level n, and analyse these.

Solution. There are many methods. The common idea is to switch between degrees d and d' in some way. For (a), if vertices have random degree with expected degree $1/p$, then $p_c = p$ since the percolation is a Bienayme-Galton-Watson tree.

For (b) you need to carefully switch between levels with degree d and d' so that $\theta(p) > 0$. For example, if $p = 1/d$ then take $d' = d + 1$, and have very fey levels with degree d'. The analysis is rather delicate.