SOLUTIONS

Correlation and concentration:

1. One can build quite small examples of this. Think of $B_1 \cup B_2$ as $E_{11} \cup E_{10} \cup E_{01}$ where $E_{11} = B_1 \cap B_2$, $E_{10} = B_1 \cap B_2^c$ and $E_{01} = B_1^c \cap B_2$. Then we want A to be likely on $E_{11} \cup E_{10}$ and on $E_{11} \cup E_{01}$ but unlikely on $E_{11} \cup E_{01} \cup E_{10}$, so we make A very likely on E_{11} and unlikely on E_{01} and E_{10}.

   So e.g., a 4-element probability space $\{x, y, z, w\}$ with $B_1 = \{x \cup y\}$, $B_2 = \{x \cup z\}$ and $A = \{x \cup w\}$, with the probability measure being, say, $p_x = p_y = p_z = 0.3$, $p_w = 0.1$.

2. (i) The intersection of two or more up-sets is an up-set: if $A \in \bigcap_i \mathcal{U}_i$ and $A \subseteq B \subseteq \Omega$, then for each i we have $A \in \mathcal{U}_i$ and hence $B \in \mathcal{U}_i$, so $B \in \bigcap_i \mathcal{U}_i$. So by Harris's Lemma we have

   $$\mathbb{P}_p(\mathcal{U}_1 \cap \cdots \cap \mathcal{U}_k) \geqslant \mathbb{P}_p(\mathcal{U}_1 \cap \cdots \cap \mathcal{U}_{k-1})\mathbb{P}_p(\mathcal{U}_k),$$

   giving the result by induction.

   (ii) $\mathcal{D}^c$ is an up-set, so $\mathbb{P}_p(\mathcal{U}_1 \mid \mathcal{D}^c) \geqslant \mathbb{P}_p(\mathcal{U}_1)$ by Harris's Lemma. Hence

   $$\mathbb{P}_p(\mathcal{U}_1 \cap \mathcal{D}) = \mathbb{P}_p(\mathcal{U}_1) - \mathbb{P}_p(\mathcal{U}_1 \cap \mathcal{D}^c) \leqslant \mathbb{P}_p(\mathcal{U}_1) - \mathbb{P}_p(\mathcal{U}_1)\mathbb{P}_p(\mathcal{D}^c) = \mathbb{P}_p(\mathcal{U}_1)\mathbb{P}_p(\mathcal{D}),$$

   giving the result.

   (iii) We have

   $$\begin{aligned}\mathbb{P}_p(\mathcal{U}_1 \cap \mathcal{U}_2 \cap \mathcal{D}) = \mathbb{P}_p(\mathcal{U}_1 \cap \mathcal{U}_2) - \mathbb{P}_p(\mathcal{U}_1 \cap \mathcal{U}_2 \cap \mathcal{D}^c) &= \mathbb{P}_p(\mathcal{U}_1)\mathbb{P}_p(\mathcal{U}_2) - \mathbb{P}_p(\mathcal{U}_1 \cap \mathcal{U}_2 \cap \mathcal{D}^c) \\ &\leqslant \mathbb{P}_p(\mathcal{U}_1)\mathbb{P}_p(\mathcal{U}_2) - \mathbb{P}_p(\mathcal{U}_1)\mathbb{P}_p(\mathcal{U}_2 \cap \mathcal{D}^c)\end{aligned}$$

   using independence and Harris's Lemma. The last formula is just $\mathbb{P}_p(\mathcal{U}_1)\mathbb{P}_p(\mathcal{U}_2 \cap \mathcal{D})$, so dividing through by $\mathbb{P}_p(\mathcal{U}_2 \cap \mathcal{D})$ gives the result.

   (iv) Without independence it need not always hold. There are examples even with $|X| = 2$: taking $p = 1/2$, so our probability space consists of all four subsets of $\{1, 2\}$ with all four equally likely, let $\mathcal{U}_1 = \{\{1\}, \{1, 2\}\}$ be the event '$1 \in X$', let $\mathcal{U}_2 = \{\{1\}, \{2\}, \{1, 2\}\}$ be the event '$1 \in X$ or $2 \in X$', and let $\mathcal{D} = \{\emptyset, \{1\}\}$ be the event '$2 \notin X$'.

3. The nice solution is to construct a coupling. Let $p_1 < p_2$ be probabilities, and construct random sets Y and Z as follows: for each element of the ground-set X, include it into both Y and Z with probability p_1, into just Z with probability $p_2 - p_1$, and into neither with probability $1-p_2$. Make this choice independently for each $i \in X$. Then Y has the distribution of the random subset X_{p_1}, and Z that of X_{p_2}. So $f(p_1) = \mathbb{P}(Y \in \mathcal{U})$ and $f(p_2) = \mathbb{P}(Z \in \mathcal{U})$. But $Y \subseteq Z$ always holds, so whenever $Y \in \mathcal{U}$ then $Z \in \mathcal{U}$. So $f(p_2) \geqslant f(p_1)$.

   $f(p) = \sum_{A \in \mathcal{U}} p^{|A|}(1-p)^{n-|A|}$ is a polynomial, so it is infinitely differentiable.

4. We assume $0 < p < 1$ to make this non-trivial. Looking at the proof in lectures, if equality holds then in the notation of that proof we must have $a_0 = a_1$ *or* $b_0 = b_1$. This in turn means $\mathcal{A}_0 = \mathcal{A}_1$ or $\mathcal{B}_0 = \mathcal{B}_1$. I.e., one of the up-sets $\mathcal{A}$ or $\mathcal{B}$ does not depend on coordinate n. (Meaning whether or not $X_p \in \mathcal{A}$, for example, depends only on which elements of $\{1, 2, \ldots, n-1\}$ are in X_p, not whether n is.)

But the same applies replacing n by any other coordinate. So we find that there are disjoint sets S, $T \subset [n]$ of coordinates such that $\mathcal{A}$ only depends on coordinates in S, and $\mathcal{B}$ only depends on coordinates in T. The existence of such sets characterizes independence for up-sets, since $\mathcal{A}$ and $\mathcal{B}$ are clearly independent in this case.

5. We apply Janson's inequality to X, the number of triangles in $G(n,p)$. Let A_i be the event that the ith possible triangle is present, for $1 \leqslant i \leqslant N = \binom{n}{3}$. Clearly $\mathbb{P}(A_i) = p^3$, so $\mu = Np^3 = \Theta(n^3)$ (since $p > 0$ is constant).

   For Δ, note that $i \sim j$ if and only if the corresponding triangles share at least one edge, which happens only when they have exactly two common vertices. There are $\Theta(n^4)$ such pairs of triangles, so $\Delta = O(n^4p^5) = O(n^4)$. Hence both μ and μ^2/Δ are at least constants times n^2 (the μ^2/Δ term dominates), and by the extended Janson inequality (Corollary 6.5) we have $\mathbb{P}(X = 0) \leqslant \exp(-cn^2)$ for some constant $c > 0$.

   For the next part, on might try arguing as follows: $\mathbb{P}(X = 0) = \mathbb{P}(\cap A_i^c) \geqslant \prod_i \mathbb{P}(A_i^c)$ by Harris's Lemma (and induction), which applies to two down-sets just as well as to two up-sets (argue as in Q2(ii), twice). This is just $(1-p^3)^N = \exp(-\Theta(n^3))$. This is a valid lower bound, but it turns out not a good one: the probability that there are no triangles is at least the probability that $G(n,p)$ contains no edges at all, namely $(1-p)^{\binom{n}{2}} = \exp(-\Theta(n^2))$. (All non-zero probabilities in $G(n,p)$ are at least $\exp(-\Theta(n^2))$.)

   Turning to triangle-free subsets: let $k = \lceil A \log n \rceil$. We want to show that whp there is no triangle-free subset of size k (then there is no larger one, of course). By the first part, the probability that a given set V of size k is triangle-free is at most e^{-ck^2}. There are $\binom{n}{k} \leqslant n^k$ subsets to consider, so by the union bound the probability that one or more is triangle-free is at most $n^k e^{-ck^2} = \exp\big(k(\log n - ck)\big)$. This tends to 0 if we choose A larger than $1/c$.

   The answer to the last part is (should be!) yes: a first moment argument would say that for A smaller than $1/c'$, the expected number of such sets is large (which does *not* prove that whp there is one). In fact, we know the answer, since by the results of the last section of the course, whp there is an independent set of size at least $B \log n$ for some B. An independent set is of course triangle-free.

6. With X the number of copies of H in $G(n,p)$, from the lecture notes we have that $\mu = \mathbb{E}[X] = \Theta(n^v p^e)$ where $v = |H|$ and $e = e(H)$. (Actually, we have an exact formula, but this is not needed.) Thus $\mu = \Theta(n^{v-\alpha e}) = \Theta(n^c)$ where $c > 0$ by the given condition on α.

   Turning to Δ, this is exactly the same quantity as calculated in the proof of Theorem 2.5. Arguing exactly as there, $\Delta = \Delta(n)$ is a finite sum (a sum where the number of terms does not vary with n) of terms each of which is $\Theta(\mu^2/(n^r p^s))$ for some (different for each term) values r, s satisfying $s/r \leqslant e/v$. (This is because r and s are possible numbers of vertices and edges in a subgraph of H, arising as an intersection of two copies of H.) Consider one such term. We have $s < re/v$, so $n^r p^s = n^{r-\alpha s} = n^{c'}$ where $c' = r - \alpha s > r - \alpha re/v > 0$ by the condition on α. So $\Delta = \Delta(n) = \sum_{i=1}^k \Delta_i(n)$ where $\Delta_i(n) = \Theta(\mu^2/n^{c_i'})$ with different c_i for each term. It follows that $\Delta = O(\mu^2/n^{c'})$ for some $c' > 0$ (the smallest of the c_i').

   Now we apply Janson's inequality in the form of Corollary 6.5. The key quantity $\min\{\mu/2, \mu^2/(2\Delta)\}$ is at least $\Theta(n^x)$ for some $x = \min\{c, c'\} > 0$ so (taking any $0 < \beta < x$) it is at least n^β for n sufficiently large. Hence $\mathbb{P}(X = 0) \leqslant \exp(-n^\beta)$ for n sufficiently large.

   [The key point is to show that the key quantity is of order a positive power of n; converting to at least a power of n for n large is just a small extra step.]

Extra questions:

7. *Connectedness of* $G(n,p)$ (i) Y_2 is the number of components of $G(n,p)$ consisting of a single edge. For a given edge $e = xy$, it is a component if and only if (a) e is present in $G(n,p)$, and (b) none of the $2(n-2)$ other possible edges between $\{x,y\}$ and the remaining vertices are present in $G(n,p)$. If A_e is the event that e is such a component, we have that Y_2 is the number of events A_e that hold, so

$$\mathbb{E}[Y_2] = \sum_e \mathbb{P}(A_e) = \binom{n}{2} p(1-p)^{2(n-2)} \sim \frac{n^2}{2} p(1-p)^{2n}.$$

Using $1 - p \leqslant e^{-p}$ this gives

$$\mathbb{E}[Y_2] \leqslant n^2 p e^{-2np} = n(\log n + c)e^{-2\log n - 2c} = \frac{(\log n + c)}{n} e^{-2c} \to 0$$

as $n \to \infty$.

(ii) Fix a set V of r vertices, say $\{1, 2, \ldots, r\}$. Let $\pi = \mathbb{P}(E)$ be the probability of the event E that these specific r vertices form a component of $G(n,p)$. Then by linearity of expectation (and symmetry), we have $\mathbb{E}[Y_r] = \binom{n}{r}\pi$.

Now E holds exactly when (a) the subgraph of $G(n,p)$ induced by V is connected, and so contains a spanning tree, and (b) there are no edges from V to $[n] \setminus V$. For (a), there are r^{r-2} possible trees on V, each of which is included in $G(n,p)$ with probability p^{r-1}. So by the union bound the probability π_a of (a) is at most $r^{r-2}p^{r-1}$. (This is only an upper bound, since we might have a component which is not a tree, and so has several spanning trees inside it.) The probability π_b of (b) is $(1-p)^{r(n-r)}$. The events involve disjoint sets of edges, so $\pi = \pi_a \pi_b \leqslant r^{r-2}p^{r-1}(1-p)^{r(n-r)}$, giving the claimed bound.

Following the plan in the hint, from the bound we have just calculated we have

$$\mathbb{E}Y_r \leqslant \left(\frac{en}{r}\right)^r r^{r-2} p^{r-1} e^{-pr(n-r)}$$

for every n and r, which we can write as

$$\mathbb{E}Y_r \leqslant r^{-2} p^{-1} x_{n,r}^r$$

where

$$x_{n,r} = \frac{en}{r} rpe^{-p(n-r)} = enpe^{-p(n-r)}.$$

We only consider terms where $r \leqslant n/2$, so $n - r \geqslant n/2$ and then

$$x_{n,r} \leqslant x_n = enpe^{-np/2}.$$

From the bounds on np in the hint, for n large we have

$$x_n \leqslant 2e(\log n)e^{-3\log n/8} = 2e(\log n)n^{-3/8} = o(n^{-1/3}).$$

Now

$$\mathbb{E}\sum_{r=3}^{\lfloor n/2 \rfloor} Y_r = \sum_{r=3}^{\lfloor n/2 \rfloor} \mathbb{E}Y_r \leqslant \sum_{r=3}^{\lfloor n/2 \rfloor} r^{-2}p^{-1}x_n^r \leqslant \sum_{r=3}^{\lfloor n/2 \rfloor} nx_n^r,$$

again using a bound in the hint. Finally, this is at most

$$n\sum_{r=3}^{\infty} x_n^r = \frac{nx_n^3}{1 - x_n} = O(nx_n^3),$$

since $x_n \to 0$ (just $x_n \leqslant 0.999$ for large n is enough for this). Since $x_n = o(n^{-1/3})$ the final bound is $o(1)$.

(iii) Suppose it is not the case that all non-isolated vertices are in a single component. Then there are two components with at least two vertices. Considering the smaller one, there is a component with between 2 and $n/2$ vertices. The probability of this is (by Markov) at most the expected number of such components, i.e., $\mathbb{E}\sum_{r=2}^{\lfloor n/2 \rfloor} Y_r$, which by (i) and (ii) is $o(1)$.

(iv) Let $f(n)$ tend to infinity sufficiently slowly that the result in part (iii) holds with $p = (\log n \pm f(n))/n$. (The question allows us to assume such an $f(n)$ exists.) Taking $p = p_+ = (\log n + f(n))/n$, by Sheet 2 Q3, whp $G(n,p)$ has no isolated vertices, so by part (iii) whp $G(n,p)$ is connected. Taking $p = p_- = (\log n - f(n))/n$, by Sheet 2 Q3, whp $G(n,p)$ has at least one isolated vertex, so (trivially) whp $G(n,p)$ is not connected. This establishes the threshold result.

More specifically, let $p^* = p^*(n) = (\log n)/n$. Suppose that $p = p(n)$ satisfies $p/p^* \to \infty$. Then $p \geqslant p_+$ for n sufficiently large, so

$$\mathbb{P}(G(n,p) \text{ is connected}) \geqslant \mathbb{P}(G(n,p_+) \text{ is connected}) \to 1.$$

Suppose that $p = p(n)$ satisfies $p = o(p^*)$. Then $p \leqslant p_-$ for n sufficiently large, so

$$\mathbb{P}(G(n,p) \text{ is connected}) \leqslant \mathbb{P}(G(n,p_-) \text{ is connected}) \to 0.$$

8. *Buffon's needle.* Not really expecting something extremely formal here!

For the first part, we have a straight needle of length $\ell < 1$. It can't cross more than one line (or the same line twice) so the number N of crossings is either 0 or 1, and $e(\ell) = 0\mathbb{P}(N = 0) + 1\mathbb{P}(N = 1) = \mathbb{P}(N = 1) = p(\ell)$.

For the next part, one solution is to consider a needle of length $\ell_1 + \ell_2$ as made up of two needles of lengths ℓ_1 and ℓ_2 stuck together. If we drop the whole needle randomly, each part considered on its own falls randomly (but not independently, of the other one, of course). Writing N_i for the number of lines the ith part crosses, we have $N = N_1 + N_2$ so by linearity of expectation (which does not need independence) $\mathbb{E}[N] = \mathbb{E}[N_1] + \mathbb{E}[N_2]$, i.e., $e(\ell_1 + \ell_2) = e(\ell_1) + e(\ell_2)$. It follows by first year maths (using continuity) that $e(\ell) = c\ell$ for some constant c. (This part does not require $\ell < 1$.)

Finally: for a given length, the expectation does not depend on the shape if we allow curved needles! To see this, suppose we have a curved needle of some length ℓ. We'll assume the curve is reasonably smooth. Then our needle can be well approximated by a sequence of n straight line segments each of length ℓ/n. More precisely, we can find such an approximation for each n with the shape converging to that of our needle, and hence the expected number of crossings converging.

Consider such a piecewise linear needle. As above we have $\mathbb{E}[N] = \sum_{i=1}^{n} \mathbb{E}[N_i]$ where N_i is the number of lines the ith segment crosses. From above, $\mathbb{E}[N_i] = p(\ell/n) = e(\ell/n) = c\ell/n$ if $n > \ell$, so $\mathbb{E}[N] = c\ell$ for such a piecewise linear needle, not just for straight ones.

Thus, for two curved needles of length ℓ, we can find arbitrarily close linear approximations having the same expected number of crossings, so the expected number of crossings for the two needles is exactly equal.

What's an easy shape? A circle of radius 1/2, so diameter 1. This always has exactly two crossings, so the expectation is 2 and $e(\pi) = 2$. Thus $c = 2/\pi$.

Bonus question

9. It's easiest to separate out the effect of the different coordinates, considering $f(\mathbf{p}) = \mathbb{P}(X_{\mathbf{p}} \in \mathcal{U})$ where $\mathbf{p} = (p_1, \ldots, p_n)$, and $X_{\mathbf{p}}$ is the random subset of our ground-set $X = \{1, 2, \ldots, n\}$ obtained by including each element i with probability p_i, independently for the different i.

   A compact way to write this is to use the notation of the proof of Harris's Lemma. In that notation we have (as in that proof)

$$f(\mathbf{p}) = (1 - p_n)\mathbb{P}_{\mathbf{p}'}(\mathcal{U}_0) + p_n\mathbb{P}_{\mathbf{p}'}(\mathcal{U}_1)$$

   where $\mathbf{p}' = (p_1, \ldots, p_{n-1})$. The only dependence on p_n on the right is the obvious one, so

$$\frac{\partial}{\partial p_n} f(\mathbf{p}) = -\mathbb{P}_{\mathbf{p}'}(\mathcal{U}_0) + \mathbb{P}_{\mathbf{p}'}(\mathcal{U}_1) = \mathbb{P}_{\mathbf{p}'}(\mathcal{U}_1 \setminus \mathcal{U}_0),$$

   since $\mathcal{U}_0 \subseteq \mathcal{U}_1$. Evaluating this partial derivative at the point $\mathbf{p} = (p, p, \ldots, p)$ we obtain

$$\mathbb{P}_p(\mathcal{U}_1 \setminus \mathcal{U}_0) = I_p(n, \mathcal{U}).$$

   To see the last equality, let $X'_p = X_p \setminus \{n\} = (X \setminus \{n\})_p$. As $\mathcal{U}$ is an up-set, we can't have $X_p \setminus \{n\} \in \mathcal{U}$ and $X_p \cup \{n\} \notin \mathcal{U}$, so exactly one of $X_p \setminus \{n\}$ and $X_p \cup \{n\}$ is in $\mathcal{U}$ if and only if $X_p \setminus \{n\} \notin \mathcal{U}$ and $X_p \cup \{n\} \in \mathcal{U}$. Whether this holds doesn't depend on whether $n \in X_p$. The condition is exactly that $X'_p \notin \mathcal{U}_0$ and $X'_p \in \mathcal{U}_1$, i.e., $X'_p \in \mathcal{U}_1 \setminus \mathcal{U}_0$.

   Now, using first year calculus for the first step,

$$\frac{\mathrm{d}}{\mathrm{d}p} f(p) = \sum_{i=1}^{n} \frac{\partial}{\partial p_n} f(\mathbf{p})\bigg|_{\mathbf{p}=(p,p,\ldots,p)} = \sum_{i=1}^{n} I_p(i, \mathcal{U}),$$

   as required.