Technische Universität München
Zentrum Mathematik
Random Graphs (MA 5208) Summer 2012
Prof. Dr. Marc Noy | Prof. Dr. Anusch Taraz | M.Sc. Peter Heinig

Sheet 4

Problem 4.1 *Composing the binomial distribution with a uniform distribution.*
Let X have the binomial distribution $\mathrm{bin}(n, U)$, where U is uniform on $(0, 1)$. Show that X is uniformly distributed on $\{0, 1, \ldots, n\}$.

A solution. We recall that if X and Y are random variables over a probability space $(\Omega, \mathcal{F}, \mathsf{Pr})$ with $\sum_{x\in \mathrm{im}(X)} |x| \cdot \mathsf{Pr}[X = x] < \infty$, then for every $\mathcal{A} \in \mathcal{F}$ with $\mathsf{Pr}[\mathcal{A}] > 0$ the *conditional expectation of* X *given* $\mathcal{A}$ is defined as

$$\mathsf{Ex}[X \mid \mathcal{A}] := \tfrac{\mathsf{Ex}[\mathbf{1}_{\mathcal{A}} \cdot X]}{\mathsf{Pr}[\mathcal{A}]} = \sum_{x\in \mathrm{im}(X)} x \cdot \mathsf{Pr}[X = x \mid \mathcal{A}] \quad . \tag{1}$$

where $\mathbf{1}_{\mathcal{A}}$ denotes the indicator variable of $\mathcal{A}$. The function

$$\mathrm{im}(Y) \ni y \mapsto \mathsf{Ex}[X \mid Y = y]$$

is called *conditional expectation given* Y and denoted by $\mathsf{Ex}[X \mid Y]$. (These are specializations of the general definition where one conditions on an arbitrary sub-σ-algebra.)

The following simple identity can be a useful tool for computing the expectation $\mathsf{Ex}[X]$ of a random variable X which depends on other random events, allowing us to treat these outcomes one by one:

Lemma 1 *Let X and Y be random variables, with X integrable. Then* $\mathsf{Ex}[\mathsf{Ex}[X \mid Y]] = \mathsf{Ex}[X]$. □

Lemma 2 (pgf of the discrete uniform distribution) . *With the abbreviation* $[n) := \{0, 1, \ldots, n-1\}$, *let* $\mathrm{Uni}([n))$ *denote be uniform distributon on* $\{0, 1, \ldots, n-1\}$. *Then* $\mathrm{pgf}(\mathrm{Uni}([n)))(z) = \frac{1}{n}\,\frac{z^n-1}{z-1}$.

Proof of Lemma 2. By definition of a pgf we have $\mathrm{pgf}(U^{[n)})\ \mathsf{Ex}(z^{U^{[n)}}) = \sum_{k\in \mathrm{im}(U^{[n)})} \mathsf{Pr}[U^{[n)} = k]z^k$ $= \sum_{0\leq k\leq n-1} \mathsf{Pr}[U^{[n)} = k]z^k = \frac{1}{n}\sum_{0\leq k\leq n} z^k = \frac{1}{n}\frac{z^n-1}{z-1}$, completing the proof. □

With these preparations we turn to Problem 4.1: we have $\mathsf{Ex}[u^X] = \mathsf{Ex}[\mathsf{Ex}[u^X \mid U]] =$ (definition of expectation of $(0, 1)$-valued a random variable) $= \int_0^1 \mathsf{Ex}[u^X \mid U = p]\mathrm{d}p = \int_0^1 \sum_{0\leq k\leq n} \mathsf{Pr}[X = k \mid U = p]\ u^k\ \mathrm{d}p = \int_0^1 \sum_{0\leq k\leq n} \mathsf{Pr}[\mathrm{binom}(n, p) = k]\ u^k\ \mathrm{d}p = \int_0^1 \mathrm{pgf}(\mathrm{binom}(n, p))(u)\mathrm{d}p =$ (lecture of 30 may 2012) $= \int_0^1 ((1-p) + pu)^n \mathrm{d}p = \int_0^1 (1 + p(u-1))^n \mathrm{d}p = \frac{1}{n+1}\frac{1}{u-1}(1 + p(u-1))^{n+1}\Big|_{p=0}^{p=1} =$ $\frac{1}{n+1}\frac{1}{u-1}u^{n+1} - \frac{1}{n+1}\frac{1}{u-1} = \frac{1}{n+1}\frac{u^{n+1}-1}{u-1}$. By Lemma 2 this is $\mathrm{pgf}(\mathrm{Uni}([n+1)))(u)$. Since equality of pgf implies equality of distributions, this proves the claim of Problem 4.1.

Problem 4.2 *Generating function for size of the n-th generation.*
Let the offspring distribution of a branching process be $p_k = (1-p)p^k$ for $k \geq 0$ and $0 \leq p < 1$. Show that the generating function $G_n(s)$ of the family size Z_n at generation n (assuming $Z_0 = 1$) is

equal to

$$G_n(u) = \begin{cases} \frac{n-(n-1)u}{n+1-nu} & \text{if } p = q = 1/2 \quad , \\ \frac{q[p^n - q^n - pu(p^{n-1}-q^{n-1})]}{p^{n+1}-q^{n+1}-pu(p^n-q^n)} & \text{if } p \neq q \quad . \end{cases} \tag{2}$$

Deduce that the probability of ultimate extinction is

$$\lim_{n\to\infty} \mathsf{Pr}[Z_n = 0] = \begin{cases} 1 & \text{if } p \leq q \quad , \\ q/p & \text{if } q < p \quad . \end{cases} \tag{3}$$

A solution. Let us first note the trival erratum that in the original problem statement we should have written '$0 \leq p < 1$' instead of '$0 \leq p \leq 1$' since for $p = 1$, every $k \in \mathbb{Z}_{\geq 0}$ is assigned 0, so we would not have a probability measure. (And, yes, while we are at it, the degenerate case $p = 0$ is not quite clear either, due to the appearance of 0^0; this is why Definition 4 below treats this case separately.)

We recall that it was mentioned in the lecture of 30 may 2012 (first theorem in the section on branching processes) that for the basic Galton–Watson process, the n-fold iterated composition of the offspring distribution equals the pgf of Z_n, the size of the n-th generation:

Lemma 3 *Let (Z_n) denote the basic Galton–Watson process with $Z_0 = 1$ and the offspring given by a fixed $\mathbb{Z}_{\geq 0}$-valued random variable* N. *Then $G_{Z_n}(u) = G_{\mathrm{N}}^{\circ n}(u)$, for every $n \geq 1$.*

Proof of Lemma 3. For $n = 1$, we have $G_{Z_1}(u) =$ (by definition of the process, Z_1 is a copy of N) $= G_{\mathrm{N}}(u) = G_{\mathrm{N}}^{\circ 1}(u)$, so the claim is true. Now assume that $n > 1$ and that $G_{Z_{n-1}}(u) = G_{\mathrm{N}}^{\circ(n-1)}(u)$. Then

$$\begin{aligned}
G_{Z_n}(u) = \mathsf{Ex}[u^{Z_n}] &= \mathsf{Ex}[\mathsf{Ex}[u^{Z_n} \mid Z_{n-1}]] \\
&= \sum_{a\geq 0} \mathsf{Ex}[u^{Z_n} \mid Z_{n-1} = a]\, \mathsf{Pr}[Z_{n-1} = a] \\
\text{(by Lemma 10, with } \mathrm{N}_1,\dots,\mathrm{N}_a \text{ denoting independent copies of N)} &= \sum_{a\geq 0} \mathsf{Pr}[Z_{n-1} = a] \sum_{b\geq 0} u^b\, \mathsf{Pr}[\mathrm{N}_1 + \cdots + \mathrm{N}_a = b] \\
&= \sum_{a\geq 0} \mathsf{Pr}[Z_{n-1} = a]\, \mathsf{Ex}[u^{\mathrm{N}_1+\cdots+\mathrm{N}_a}] = \sum_{a\geq 0} \mathsf{Pr}[Z_{n-1} = a]\, \mathsf{Ex}[u^{\mathrm{N}_1} \cdots u^{\mathrm{N}_a}] \\
\text{(}\mathsf{Ex}[\cdot] \text{ multiplicative for independent random variables)} &= \sum_{a\geq 0} \mathsf{Pr}[Z_{n-1} = a]\, \mathsf{Ex}[u^{\mathrm{N}_1}] \cdots \mathsf{Ex}[u^{\mathrm{N}_a}] = \sum_{a\geq 0} \mathsf{Pr}[Z_{n-1} = a] (\mathsf{Ex}[u^{\mathrm{N}}])^a \\
&= \sum_{a\geq 0} \mathsf{Pr}[Z_{n-1} = a] (G_{\mathrm{N}}(u))^a = (\sum_{a\geq 0} \mathsf{Pr}[Z_{n-1} = a] z^a)|_{z=G_{\mathrm{N}}(u)} \\
&= (\mathsf{Ex}[z^{Z_{n-1}}])|_{z=G_{\mathrm{N}}(u)} = G_{Z_{n-1}}(z)\, |_{z=G_{\mathrm{N}}(u)} \\
\text{(inductive hypothesis)} &= G_{\mathrm{N}}^{\circ(n-1)}(z)\, |_{z=G_{\mathrm{N}}(u)} = G_{\mathrm{N}}^{\circ(n-1)}(G_{\mathrm{N}}(u)) = G_{\mathrm{N}}^{\circ n}(u) \quad ,
\end{aligned} \tag{4}$$

which completes the proof of Lemma 3. □

We abbreviate $G_n(u) := G_{\mathrm{N}}^{\circ n}(u)$. Problem 4.2 tells us that if the offspring distribution is the geometric distribution[1] geo(p), then using Lemma 3 we can find an expression for $G_{Z_n}(u)$ as a rational

[1]Caution: there are two variants of the geometric distribution, according to whether 0 is in the support (for the distribution used in the problem it is). The pgf stated in the lecture of 30 may 2012 was for the other variant and is not compatible with the present problem.

function.

Definition 4 (geometric distribution with 0 in the support) . *For* $p = 0$ *define* $\mathrm{geo}(p)$ *as the probability distribution with mass function* $0 \mapsto 1$ *and* $k \mapsto 0$ *for every* $k \geq 1$*. For every* $0 < p < 1$ *let* $\mathrm{geo}(p)$ *denote the discrete probability distribution with support* $\mathbb{Z}_{\geq 0}$ *and mass function* $k \mapsto (1-p)p^k$.

Lemma 5 (pgf of the geometric distribution with 0 in the support) . *For every* $0 \leq p \leq 1$ *we have* $\mathrm{pgf}(\mathrm{geo}(p))(z) = G_{\mathrm{geo}(p)}(z) = \frac{1-p}{1-pz}$.

Proof of Lemma 5. For $p = 0$ indeed $\mathrm{pgf}(\mathrm{geo}(p)) = 1 = \frac{1-0}{1-0z}$. For $0 < p < 1$ we can calculate $\mathrm{pgf}(\mathrm{geo}(p))(z) = \mathsf{Ex}[z^{\mathrm{geo}(p)}] = \sum_{k\geq 0} \mathsf{Pr}[\mathrm{geo}(p) = k]\ z^k = \sum_{k\geq 0}(1-p)p^k z^k = (1-p)\sum_{k\geq 0}(pz)^k = (1-p)\sum_{k\geq 0}(pz)^k = \frac{1-p}{1-pz}$. □

After these preparations we turn to (2):

Case that $p = \frac{1}{2}$. Induction on n. Since $G_1(u) =$ (definition) $= G_{\mathrm{geo}(p)}(u) =$ (by Lemma 5) $= \frac{1-\frac{1}{2}}{1-\frac{1}{2}u}$ $= \frac{1}{2-u} = \frac{n-(n-1)u}{n+1-nu}\big|_{n=1}$, hence (2) is true for $n = 1$. Now assume that $n > 1$ and that the case $p = \frac{1}{2}$ of (2) holds for $n-1$. We have $G_n(u) =$ (definition) $= G_{\mathrm{geo}(p)}(G_{n-1}(u)) =$ (inductive assumption) $= G_{\mathrm{geo}(p)}(\frac{n-1-(n-2)u}{n-(n-1)u}) =$ (Lemma 5) $= \frac{1}{2-\frac{n-1-(n-2)u}{n-(n-1)u}} = \frac{n-(n-1)u}{2(n-(n-1)u)-(n-1-(n-2)u)} = \frac{n-(n-1)u}{n+1-nu}$, which proves (2) in the case $p = \frac{1}{2}$.

Case that $p \neq \frac{1}{2}$. Induction on n. Since we have $\left.\frac{q[p^n-q^n-pu(p^{n-1}-q^{n-1})]}{p^{n+1}-q^{n+1}-pu(p^n-q^n)}\right|_{n=1} = \frac{q(p-q)}{p^2-q^2-pu(p-q)} = \frac{q}{p+q-pu} = \frac{1-p}{1-pu} =$ (by Lemma 4) $= G_{\mathrm{geo}(p)}(u)$, the expression claimed in the case $p \neq q$ in (2) is true for $n = 1$. Now assume that $n > 1$ and that the case $p \neq \frac{1}{2}$ holds for (2). Then we can calculate as follows: $G_n(u) =$ (definition) $= G_{\mathrm{geo}(p)}(G_{n-1}(u)) =$ (inductive assumption) $= G_{\mathrm{geo}(p)}(\frac{q[p^{n-1}-q^{n-1}-pu(p^{n-2}-q^{n-2})]}{p^n-q^n-pu(p^{n-1}-q^{n-1})}) =$ (by Lemma 5 with the numerator $1-p$ denoted q) $=$ $\frac{q}{1-p\cdot\frac{q(p^{n-1}-q^{n-1}-pu(p^{n-2}-q^{n-2}))}{p^n-q^n-pu(p^{n-1}-q^{n-1})}} = \frac{q[p^n-q^n-pu(p^{n-1}-q^{n-1})]}{p^n-q^n-pu(p^{n-1}-q^{n-1})-p\cdot q(p^{n-1}-q^{n-1}-pu(p^{n-2}-q^{n-2}))}$. Twice applying the identity

$$pq(p^a-q^a) = (p^{a+1}-q^{a+1}) - (p^{a+2}-q^{a+2}) \quad , \tag{5}$$

valid because of $p + q = 1$, we can transform the denominator of the fraction as follows: $p^n - q^n - pu(p^{n-1}-q^{n-1}) - p \cdot q(p^{n-1}-q^{n-1}-pu(p^{n-2}-q^{n-2})) =$ (by (5)) $= p^n - q^n - pu(p^{n-1}-q^{n-1}) - (p^n - q^n) + (p^{n+1}-q^{n+1}) + pu\ pq\ (p^{n-2}-q^{n-2}) = (p^{n+1}-q^{n+1}) - pu(p^{n-1}-q^{n-1}) + pu\ pq\ (p^{n-2}-q^{n-2}) =$ (by (5)) $= p^{n+1}-q^{n+1}-pu(p^n-q^n)$. Hence the fraction we reached is the expression claimed to be $G_n(u)$ in (2). This completes the proof.

Now let Ω be some set-theoretical realization of the processes we are considering. Then

$$\{\omega \in \Omega \colon Z_t(\omega) = 0\} \subset \{\omega \in \Omega \colon Z_{t+1}(\omega) = 0\} \quad , \tag{6}$$

and (abbreviating $\{Z_t = 0\} := \{\omega \in \Omega \colon Z_t(\omega) = 0\}$),

$$\begin{aligned} \mathsf{Pr}[\{\omega \in \Omega \colon \omega \text{ ends with extinction}] &= \bigcup_{n\geq 1}\{\omega \in \Omega \colon Z_n(\omega) = 0\} \\ &= \{Z_1 = 0\} \cup (\{Z_2 = 0\} \setminus \{Z_1 = 0\}) \\ &\cup (\{Z_3 = 0\} \setminus \{Z_2 = 0\}) \cup (\{Z_4 = 0\} \setminus \{Z_3 = 0\}) \cup \cdots \quad . \end{aligned} \tag{7}$$

We can use this to calculate that

$$\begin{aligned}
\Pr[\{\omega\in\Omega\colon \omega \text{ ends with extinction}\}] &\overset{(7)}{=} \Pr\Big[\{Z_1=0\}\cup\bigcup_{n\geq 1}(\{Z_{n+1}=0\}\setminus\{Z_n=0\})\Big] \\
&= \Pr[\{Z_1=0\}] + \sum_{t\geq 2}\Pr[\{Z_t=0\}\setminus\{Z_{t-1}=0\}]
\end{aligned}$$

(since (6) implies that the sets united in (7) are not merely consecutively- but also mutually-disjoint, we can use countable additivity of a probability measure)

$$\begin{aligned}
&= \Pr[\{Z_1=0\}] + \lim_{\mathbb{Z}\ni b\to\infty}\sum_{2\leq t\leq b}\Pr[\{Z_t=0\}\setminus\{Z_{t-1}=0\}] \\
&= \Pr[\{Z_1=0\}] + \lim_{\mathbb{Z}\ni b\to\infty}\Pr[\{Z_b=0\}] - \Pr[\{Z_1=0\}] \\
&= \lim_{\mathbb{Z}\ni b\to\infty}\Pr[\{Z_b=0\}] \\
&= \lim_{\mathbb{Z}\ni b\to\infty} G_b(0)
\end{aligned}$$

(by (2))

$$= \begin{cases} \lim_{\mathbb{Z}\ni b\to\infty}\frac{b}{b+1} & \text{if } p=\frac{1}{2}\quad, \\ \lim_{\mathbb{Z}\ni b\to\infty}\frac{q(p^b-q^b)}{p^{b+1}-q^{b+1}} & \text{if } p\neq q\quad. \end{cases} \tag{8}$$

Since $\frac{q(p^b-q^b)}{p^{b+1}-q^{b+1}} = \frac{q^{b+1}-qp^b}{q^{b+1}-p^{b+1}} = \frac{1-(\frac{p}{q})^b}{1-(\frac{p}{q})^{b+1}}$, if $p<q$, then $\lim_{\mathbb{Z}\ni b\to\infty}\frac{q(p^b-q^b)}{p^{b+1}-q^{b+1}} = \frac{\lim_b 1-(\frac{p}{q})^b}{\lim_b 1-(\frac{p}{q})^{b+1}} = \frac{1}{1} = 1$.

Since $\frac{q(p^b-q^b)}{p^{b+1}-q^{b+1}} = \frac{q}{p}\cdot\frac{1-(\frac{q}{p})^b}{1-(\frac{q}{p})^{b+1}}$, if $p>q$, then $\lim_{\mathbb{Z}\ni b\to\infty}\frac{q(p^b-q^b)}{p^{b+1}-q^{b+1}} = \frac{q}{p}\cdot\frac{\lim_b 1-(\frac{q}{p})^b}{\lim_b 1-(\frac{q}{p})^{b+1}} = \frac{q}{p}$.

This completes the solution of Problem 4.2.

Problem 4.3 *Distribution of the extinction time.*
In the same situation as in the previous problem, let $T=\min\{n\colon Z_n=0\}$ be the random variable equal to the extinction time. Find $\Pr[T=n]$. For what values of p do we have $\mathsf{Ex}[T]<+\infty$?

A solution. Let Ω denote any set-theoretical realization of the branching processes in question. Then

(Se.1) $\{\omega\in\Omega\colon Z_{t-1}(\omega)=0\}\subset\{\omega\in\Omega\colon Z_t(\omega)\}$,

(Se.2) $\{\omega\in\Omega\colon T(\omega)=t\}=\{\omega\in\Omega\colon Z_t(\omega)=0\}\setminus\{\omega\in\Omega\colon Z_{t-1}(\omega)=0\}$,

hence

$$\begin{aligned}
\Pr[T=t] &= \Pr[\{\omega\in\Omega\colon Z_t(\omega)=0\}\setminus\{\omega\in\Omega\colon Z_{t-1}(\omega)=0\}] \\
&= \Pr[\{\omega\in\Omega\colon Z_t(\omega)=0\}] \\
&\quad - \Pr[\{\omega\in\Omega\colon Z_{t-1}(\omega)=0\}\cap\{\omega\in\Omega\colon Z_t(\omega)=0\}]
\end{aligned}$$

(by (Se.1))

$$= \Pr[\{\omega\in\Omega\colon Z_t(\omega)=0\}] - \Pr[\{\omega\in\Omega\colon Z_{t-1}(\omega)=0\}]$$

(by Lemma 3)

$$= G^{\circ t}_{\mathrm{geo}(p)}(0) - G^{\circ(t-1)}_{\mathrm{geo}(p)}(0)$$

(by (2))

$$\begin{aligned}
&= \begin{cases} \frac{t}{t+1}-\frac{t-1}{t} & \text{if } p=\frac{1}{2}\quad, \\ \frac{q(p^t-q^t)}{p^{t+1}-q^{t+1}} - \frac{q(p^{t-1}-q^{t-1})}{p^t-q^t} & \text{if } p\neq\frac{1}{2}\quad. \end{cases} \\
&= \begin{cases} \frac{1}{t(t+1)} & \text{if } p=\frac{1}{2}\quad, \\ \frac{p^{t-1}q^t(p^2+q^2-2pq)}{(p^{t+1}-q^{t+1})(p^t-q^t)} & \text{if } p\neq\frac{1}{2}\quad. \end{cases}
\end{aligned} \tag{9}$$

Therefore

$$\mathsf{Ex}[T] = \sum_{t\geq 1} \mathsf{Pr}[T=t]\cdot t \quad \overset{(9)}{=} \quad \begin{cases} \sum_{t\geq 1} \frac{1}{t+1} & \text{if } p=\frac{1}{2} \quad , \\ \sum_{t\geq 1} \frac{p^{t-1}q^t(p^2+q^2-2pq)}{(p^{t+1}-q^{t+1})(p^t-q^t)}t & \text{if } p\neq\frac{1}{2} \quad . \end{cases} \tag{10}$$

If $p<q$, then with

$$b_t(p) := \tfrac{p^{t-1}q^t(p^2+q^2-2pq)}{(p^{t+1}-q^{t+1})(p^t-q^t)}t \tag{11}$$

we have

$$\left|\tfrac{b_t(p)}{b_{t-1}(p)}\right| = \left|\tfrac{p^tq-q^tp}{p^{t+1}-q^{t+1}}\cdot\tfrac{t}{t-1}\right| = \tfrac{p}{q}\left|\tfrac{\frac{q}{p}(\frac{p}{q})^{t-1}-1}{(\frac{p}{q})^{t+1}-1}\right| \xrightarrow{t\to\infty} \tfrac{p}{q} < 1 \quad , \tag{12}$$

hence if $p<q$, then by the 'ratio test', $\mathsf{Ex}[T]$ converges.

If $p=\frac{1}{2}$ we have $\mathsf{Ex}[T]=\infty$ since the harmonic series diverges.

Problem 4.4 *How sizes of generations correlate with each other.*
Let Z_n be the size of the n-th generation in a branching process with $Z_0 = 1$, $\mathsf{Ex}[Z_1] = \mu$ and $\sigma^2(Z_1) > 0$. Compute $\mathsf{Ex}[Z_n|Z_m]$ and show that $\mathsf{Ex}[Z_nZ_m] = \mu^{n-m}\,\mathsf{Ex}[Z_m^2]$ for $m\leq n$. Compute $\rho(Z_m,Z_n)$, where $\rho(\cdot,\cdot)$ denotes the standard Pearson correlation coefficient. [Hint: distinguish whether $\mu=1$ or not.]

A solution. The Pearson correlation coefficient of two random variables X and Y is

$$\rho(X,Y) := \tfrac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}[X]\ \mathsf{Var}[Y]}} = \tfrac{\mathsf{Ex}[X\ Y]-\mathsf{Ex}[X]\,\mathsf{Ex}[Y]}{\sqrt{\mathsf{Var}[X]\ \mathsf{Var}[Y]}} \quad . \tag{13}$$

We therefore need to compute $\mathsf{Cov}(Z_n,Z_m)$ and $\mathsf{Var}[Z_n]$. In particular we need to compute $\mathsf{Ex}[Z_nZ_m]$. We will organize these tasks in such a way that in the end, the only quantity that we have to determine explicitly in terms of the parameters μ and $\sigma^2(Z_1)$ is $\mathsf{Var}[Z_n]$.

Lemma 6 *For every $\mathbb{Z}_{\geq 0}$-valued random variable N with pgf G_N, every $n\geq 1$, and with $\mu := \mathsf{Ex}[\mathrm{N}]$ and $\sigma^2 := \mathsf{Var}[\mathrm{N}]$,*

(Subs.1.1) $\mathrm{D}_u G_\mathrm{N}^{\circ n}(u)\big|_{u=1} = \mu^n \quad ,$

(Subs.1.2) $\mathrm{D}_u^2 G_\mathrm{N}(u)\big|_{u=1} = \mathsf{Ex}[\mathrm{N}^2] - \mu = \sigma^2+\mu^2-\mu \quad ,$

(Subs.1.3) $\mathrm{D}_u^2 G_\mathrm{N}^{\circ n}(u)\big|_{u=1} = (\sigma^2+\mu^2-\mu)\cdot\mu^{2n-2} + \mu\cdot(\mathrm{D}_u^2 G_\mathrm{N}^{\circ(n-1)}(u)\big|_{u=1}) \quad ,$

(Subs.1.4) $\mathrm{D}_u^2 G_\mathrm{N}^{\circ n}(u)\big|_{u=1} = \begin{cases} n\sigma^2 & \text{if } \mu=1 \quad , \\ \mu^{n-1}(\sigma^2+\mu^2-\mu)(\frac{1-\mu^n}{1-\mu}) & \text{if } \mu\neq 1 \quad . \end{cases}$

Proof of Lemma 6. As to (Subs.1.1) we induct on n. The case $n=1$ is obviously true and was mentioned in the lecture of 30 may 2012. Now assume $n>1$ and the (Subs.1.1) is true for $n-1$.

Now we calculate

$$
\begin{aligned}
\mathrm{D}_u G_{\mathrm{N}}^{\circ n}(u)\Big|_{u=1} &= \mathrm{D}_u G_{\mathrm{N}}(G_{\mathrm{N}}^{\circ(n-1)}(u))\Big|_{u=1} \\
\text{(chain rule for formal power series)} \quad &= (\mathrm{D}_u G_{\mathrm{N}})(G_{\mathrm{N}}^{\circ(n-1)}(u)) \cdot \mathrm{D}_u G_{\mathrm{N}}^{\circ(n-1)}(u)\Big|_{u=1} \\
\text{(evaluating is homomorphism)} \quad &= (\mathrm{D}_u G_{\mathrm{N}})(G_{\mathrm{N}}^{\circ(n-1)}(u))\Big|_{u=1} \cdot \mathrm{D}_u G_{\mathrm{N}}^{\circ(n-1)}(u)\Big|_{u=1} \\
&= (\mathrm{D}_u G_{\mathrm{N}})(G_{\mathrm{N}}^{\circ(n-1)}(1)) \cdot \mathrm{D}_u G_{\mathrm{N}}^{\circ(n-1)}(1) \\
\text{(since } \mathrm{pgf}_X(1)=1 \text{ for any random variable } X) \quad &= (\mathrm{D}_u G_{\mathrm{N}})(1) \cdot \mathrm{D}_u G_{\mathrm{N}}^{\circ(n-1)}(1) \\
\text{(inductive hypothesis)} \quad &= \mu \cdot \mu^{n-1} = \mu^n \quad , \qquad (14)
\end{aligned}
$$

completing the proof of (Subs.1.1).

As to (Subs.1.2), see lecture of 30 may 2012.

As to (Subs.1.3)

$$
\begin{aligned}
\mathrm{D}_u^2 G_{\mathrm{N}}^{\circ n}(u)\Big|_{u=1} &= \mathrm{D}_u\Bigg(\mathrm{D}_u G_{\mathrm{N}}(G_{\mathrm{N}}^{\circ(n-1)}(u))\Bigg)\Big|_{u=1} \\
\text{(chain rule)} \quad &= \mathrm{D}_u\Bigg((\mathrm{D}_u G_{\mathrm{N}})(G_{\mathrm{N}}^{\circ(n-1)}(u)) \cdot \mathrm{D}_u G_{\mathrm{N}}^{\circ(n-1)}(u)\Bigg)\Big|_{u=1} \\
\text{(product rule and chain rule)} \quad &= \Bigg((\mathrm{D}_u^2 G_{\mathrm{N}})(G_{\mathrm{N}}^{\circ(n-1)}(u)) \cdot (\mathrm{D}_u G_{\mathrm{N}}^{\circ(n-1)}(u))^2 \\
&\quad + (\mathrm{D}_u G_{\mathrm{N}})(G_{\mathrm{N}}^{\circ(n-1)}(u)) \cdot \mathrm{D}_u^2 G_{\mathrm{N}}^{\circ(n-1)}(u)\Bigg)\Big|_{u=1} \\
\text{(evaluating is homomorphism)} \quad &= (\mathrm{D}_u^2 G_{\mathrm{N}}(u))\Big|_{u=G_{\mathrm{N}}^{\circ(n-1)}(1)} \cdot (\mathrm{D}_u G_{\mathrm{N}}^{\circ(n-1)}(u)\Big|_{u=1})^2 \\
&\quad + (\mathrm{D}_u G_{\mathrm{N}}(u))\Big|_{u=G_{\mathrm{N}}^{\circ(n-1)}(1)} \cdot \mathrm{D}_u^2 G_{\mathrm{N}}^{\circ(n-1)}(u)\Big|_{u=1} \\
\left(\begin{array}{l}\text{using } G_{\mathrm{N}}^{\circ(n-1)} = G_{Z_{n-1}} \text{ by} \\ \text{Lemma 3, and } G_{Z_{n-1}}(1)=1, \\ \text{and (Subs.1.2), and (Subs.1.1)}\end{array}\right) \quad &= (\sigma^2 + \mu^2 - \mu) \cdot \mu^{2n-2} + \mu \cdot \mathrm{D}_u^2 G_{\mathrm{N}}^{\circ(n-1)}(u)\Big|_{u=1} \quad , \qquad (15)
\end{aligned}
$$

proving (Subs.1.3).

As to (Subs.1.4), we just have to solve the linear recurrence in (Subs.1.3). If $\mu = 1$, then (Subs.1.3) implies $\mathrm{D}_u^2 G_{\mathrm{N}}^{\circ n}(u)\Big|_{u=1} = (n-1)\cdot\sigma^2 + \mathrm{D}_u^2 G_{\mathrm{N}}^{\circ n}(u)\Big|_{u=1} = (\text{ by (Subs.1.2) }) = (n-1)\cdot\sigma^2 + \sigma^2 = n\cdot\sigma^2$, as claimed. If $\mu = 1$, the formula for geometric sums gives the claimed expression. □

In the following, whereas (1) is special case of a more general fact about conditional expectations and does not use anything of the definition of the Z_n, the statements (2) and (3) do:

Lemma 7 *For $n \geq m \geq 1$ and with $\mu := \mathsf{Ex}[Z_1]$,*

(1) $\mathsf{Ex}[Z_n Z_m \mid Z_m] = Z_m \cdot \mathsf{Ex}[Z_n \mid Z_m]$,

(2) $\mathsf{Ex}[Z_n \mid Z_m] = \mu^{n-m} Z_m$,

(3) $\mathsf{Ex}[Z_n Z_m] = \mu^{n-m} \; \mathsf{Ex}[Z_m^2]$.

Proof of Lemma 13. As to (1), $\mathsf{Ex}[Z_n Z_m \mid Z_m](y) = \mathsf{Ex}[Z_n Z_m \mid Z_m = y] =$ (by the definition 1) $= \frac{\mathsf{Ex}[Z_n Z_m \cdot \mathbf{1}_{\{Z_m = y\}}]}{\mathsf{Pr}[Z_m = y]} = \frac{\mathsf{Ex}[Z_n \cdot y \cdot \mathbf{1}_{\{Z_m = y\}}]}{\mathsf{Pr}[Z_m = y]} = y \cdot \frac{\mathsf{Ex}[Z_n \cdot \mathbf{1}_{\{Z_m = y\}}]}{\mathsf{Pr}[Z_m = y]} = y \cdot \mathsf{Ex}[Z_n \mid Z_m = y] = (Z_m \cdot \mathsf{Ex}[Z_n \mid Z_m])(y)$, for every $y \in \mathrm{im}(Z_m)$, which is what (1) says.

As to (2), for every $y \in \mathbb{Z}_{\geq 0}$ we have $\mathsf{Ex}[Z_n \mid Z_m](y) = \mathsf{Ex}[Z_n \mid Z_m = y] = \sum_{x \in \mathbb{Z}_{\geq 0}} x \cdot \mathsf{Pr}[Z_n = x \mid Z_m = y] =$ (by Lemma 10, and with the notation as defined there) $= \sum_{x \in \mathbb{Z}_{\geq 0}} x \, \mathsf{Pr}[Z_{n-m,1} + \cdots + Z_{n-m,x} = y]$ $= \mathsf{Ex}[Z_{n-m,1} + \cdots + Z_{n-m,y}] = \mathsf{Ex}[Z_{n-m,1}] + \cdots + \mathsf{Ex}[Z_{n-m,y}] =$ (by (Subs.1.1) with $n = 1$) $=$ $G'_{Z_{n-m,1}}(1) + \cdots + G'_{Z_{n-m,y}}(1) =$ (by definition of $Z_{n-n,i}$) $= y \cdot G'_{Z_{n-m}}(1) = y \cdot \mathrm{D}_u G_{Z_1}^{\circ(n-m)}(u)\Big|_{u=1} =$ (by (Subs.1.1)) $= y \cdot \mu^{n-m} = (\mu^{n-m} Z_m)(y)$. Hence $\mathsf{Ex}[Z_n \mid Z_m] = \mu^{n-m} Z_m$, as functions, completing the proof of (2).

As to (3), we have $\mathsf{Ex}[Z_n Z_m] =$ (by Lemma 1) $= \mathsf{Ex}[\mathsf{Ex}[Z_n Z_m \mid Z_m]] =$ (by (1)) $= \mathsf{Ex}[Z_m \cdot \mathsf{Ex}[Z_n \mid Z_m]] =$ (by (2)) $= \mathsf{Ex}[Z_m \cdot \mu^{n-m} \cdot Z_m] = \mathsf{Ex}[\mu^{n-m} \cdot Z_m^2] =$ (linearity of expectation) $= \mu^{n-m} \, \mathsf{Ex}[Z_m^2]$, proving (3).

This completes the proof of Lemma 13. □

Lemma 8 (covariance of two generation-sizes in terms of variance of the earlier one and mean of the offspring distribution) . *Let Z_n be the size of the n-th generation in the basic Galton–Watson branching process with $Z_0 = 1$ and $\mathsf{Ex}[Z_1] =: \mu$. Then, for every $n \geq m$,*

$$\mathsf{Cov}(Z_m, Z_n) = \mu^{n-m} \; \mathsf{Var}[Z_m] \quad . \tag{16}$$

Proof of Lemma 8. We can calculate

$$\begin{aligned} \mathsf{Cov}(Z_m, Z_n) &= \mathsf{Ex}[Z_m Z_n] - \mathsf{Ex}[Z_n]\,\mathsf{Ex}[Z_m] \\ \text{(by Lemma 7.(3))} \quad &= \mathsf{Ex}(Z_m^2)\; \mu^{n-m} - \mathsf{Ex}[Z_n]\,\mathsf{Ex}[Z_m] \\ \text{(setting } m := 0 \text{ in Lemma 7.(2) we find that } \mathsf{Ex}[Z_{\tilde{n}}] = \mu^{\tilde{n}}) \quad &= \mathsf{Ex}(Z_m^2)\; \mu^{n-m} - \mu^n \mu^m \\ &= \mathsf{Ex}(Z_m^2)\; \mu^{n-m} - \mu^{n-m} (\mu^m)^2 \\ &= \mu^{n-m} (\mathsf{Ex}[Z_m^2] \; - \mathsf{Ex}[Z_m]^2) \\ &= \mu^{n-m}\, \mathsf{Var}[Z_m] \quad , \end{aligned} \tag{17}$$

completing the proof of Lemma 8 □

Now with (13) and Lemma 8 we find

$$\rho(Z_m, Z_n) = \mu^{n-m} \sqrt{\frac{\mathsf{Var}[Z_m]}{\mathsf{Var}[Z_n]}} \quad , \tag{18}$$

so we will finally have expressed the correlation coefficient in terms of the given process-parameters $\mu = \mathsf{Ex}[Z_1]$ and $\sigma^2(Z_1)$, if for every $n \geq 1$ we find an expression of $\mathsf{Var}[Z_n]$ in terms of these parameters:

Lemma 9 (variance of size of the n-th generation) . *Let Z_n be the size of the n-th generation in the basic Galton–Watson branching process with $Z_0 = 1$, $\mathsf{Ex}[Z_1] = \mu$ and variance $\sigma^2(Z_1) > 0$. Then*

$$\mathsf{Var}[Z_n] = \sigma^2(Z_n) = \begin{cases} n\sigma^2(Z_1) & \text{if } \mu = 1 \quad , \\ \frac{\mu^{n-1}(1-\mu^n)}{1-\mu}\sigma^2(Z_1) & \text{if } \mu \neq 1 \quad . \end{cases} \tag{19}$$

Proof of Lemma 9. Let $G_X := \text{pgf}(X)$ and let $'$ denote differentiation. From the lecture of 30 may 2012 we know that if X is a random variable such that both $G''_X(1)$ and $G'_X(1)$ are convergent series, then $\mathsf{Ex}[X^2] = G''_X(1) + G'_X(1)$. Therefore

$$\sigma^2(Z_n) = G''_{Z_n}(1) + G'_{Z_n}(1) - (G'_{Z_n}(1))^2 \quad . \tag{20}$$

Since $G'_{Z_n}(1) = \mathrm{D}_u G_{Z_n}(u)\Big|_{u=1}$ (by Lemma 3) $= \mathrm{D}_u G^{\circ n}_{Z_1}(u)\Big|_{u=1}$ = (by Lemma 6) = $\mathsf{Ex}[Z_1]^n = \mu^n$, we find

$$\sigma^2(Z_n) = \begin{cases} \mathrm{D}^2_u G^{\circ n}_{Z_1}(u)\big|_{u=1} & \text{if } \mu = 1 \quad , \\ \mathrm{D}^2_u G^{\circ n}_{Z_1}(u)\big|_{u=1} + \mu^n - \mu^{2n} & \text{if } \mu \neq 1 \quad . \end{cases} \tag{21}$$

If $\mu = 1$, then (21) and (Subs.1.4) obviously imply the claim in (19). If $\mu \neq 1$, then again by (21) and (Subs.1.4) we get $\sigma^2(Z_2) = \mu^{n-1}(\sigma^2 + \mu^2 - \mu)(\frac{1-\mu^n}{1-\mu}) + \mu^n - \mu^{2n} = \frac{\mu^{n-1}(1-\mu^n)}{1-\mu}\sigma^2$, as claimed in (19). This completes the proof of Lemma 9.

Combining (18) and Lemma 9 we find as an answer:

$$\rho(Z_m, Z_n) = \frac{\mathrm{Cov}(Z_m, Z_n)}{\sqrt{\mathsf{Var}(Z_m)\,\mathsf{Var}(Z_n)}} = \begin{cases} \left(\mu^{n-m}\ \frac{1-\mu^m}{1-\mu^n}\right)\right)^{\frac{1}{2}} & \text{if } \mu \neq 1 \quad , \\ \left(\frac{m}{n}\right)^{\frac{1}{2}} & \text{if } \mu = 1 \quad . \end{cases} \tag{22}$$

This completes the solution of Problem 4.4.

Appendix

The following plays a role in both Problem 4.2 and Problem 4.4, so we state it here, for reference:

Lemma 10 *Let Z_n be the size of the n-th generation in the basic Galton-Watson branching process with $Z_0 = 1$ and offspring given by independent copies of a fixed $\mathbb{Z}_{\geq 0}$-valued random variable* N. *For every $g \in \mathbb{Z}_{\geq 0}$ and $x \in \mathbb{Z}_{\geq 1}$, let $Z_{g,1}, \ldots, Z_{g,x}$ be copies of Z_g such that $Z_{g,1}, \ldots, Z_{g,x}$ are independent. Then $\mathsf{Pr}[Z_n = y \mid Z_m = x] = \mathsf{Pr}[Z_{n-m,1} + \cdots + Z_{n-m,x} = y]$ for every $n \geq m \geq 1$ and every $(x, y) \in \mathbb{Z}^2_{\geq 0}$.* □