LIMITING THE BINOMIAL OPTION PRICING FORMULA TO THE BLACK-SCHOLES FORMULA

NIUSHAN GAO
DEPARTMENT OF MATHEMATICS
RYERSON UNIVERSITY

In this note, we review the famous deduction of the Black-Scholes option pricing formula from the binomial-model option pricing formula. One sees that the approximation of binomial models is very robust: the parameters p_N, u_N, d_N in the binomial model can be selected very broadly; see Section 2.4 for a discussion. In particular, the parameter p_N does not have to be around 1/2, and d_N may not have to be equal to $1/u_N$.

1. From Binomial Model to Black-Scholes Model

1.1. **Black-Scholes Model.** In the Black-Scholes model, the stock price process over a continuous-time interval $[0,T]$ follows a geometric Brownian motion

$$S_t = S_0 e^{\mu t + \sigma B_t}, \quad t \in [0,T],$$

where $(B_t)_{t\in[0,T]}$ is a standard Brownian motion. At any moment $t \in [0,T]$, the price S_t is log-normal, namely,

$$\ln S_t = \ln S_0 + \mu t + \sigma B_t \quad \sim \quad \mathrm{N}\big(\ln S_0 + \mu t, \sigma^2 t\big).$$

The intermediate performance over an infinitesimal time interval $[t,t']$ is given by

$$\ln\left(\frac{S_{t'}}{S_t}\right) = \mu(t'-t) + \sigma(B_{t'} - B_t) \quad \sim \quad \mathrm{N}\big(\mu(t'-t), \sigma^2(t'-t)\big). \tag{1.1}$$

Note that Brownian motions have continuous paths. Thus when t' is close to t, $\frac{S_{t'}-S_t}{S_t}$ is small, so that

$$\ln\left(\frac{S_{t'}}{S_t}\right) = \ln\left(1 + \frac{S_{t'}-S_t}{S_t}\right) \approx \frac{S_{t'}-S_t}{S_t},$$

where the last term is roughly the rate of return of the stock over time $[t,t']$. In what follows, we will refer to $\ln\left(\frac{S_{t'}}{S_t}\right)$ as stock return for continuous-time models.

1.2. **Stock return in two models over infinitesimal periods.**

Date: written on November 15, 2018; revised on November 22, 2018.

1.2.1. *Discretized BS model.* We cut the continuous-time interval $[0,T]$ into N infinitesimal time intervals:

$$t=0 \longrightarrow t=\tfrac{T}{N} \longrightarrow t=\tfrac{2T}{N} \longrightarrow \dots \longrightarrow t=\tfrac{(N-1)T}{N} \longrightarrow t=T$$

By (1.1), *over each infinitesimal interval* $[t,t']=[\frac{nT}{N},\frac{(n+1)T}{N}]$, *the stock return follows the normal distribution* $\mathrm{N}\Big(\frac{\mu T}{N},\frac{\sigma^2 T}{N}\Big)$. In particular, the stock has

(1) expected rate of return
$$\mathbb{E}\Big[\ln\frac{S_{t'}}{S_t}\Big]=\frac{\mu T}{N},$$
(2) variance of rate of return
$$\mathbb{V}\Big[\ln\frac{S_{t'}}{S_t}\Big]=\frac{\sigma^2 T}{N},$$
(3) probability to increase
$$\mathbb{P}(\text{stock price increases in this period}) = \mathbb{P}\Big(\ln\frac{S_{t'}}{S_t}>0\Big)=\Phi\Big(\frac{0-\mu\frac{T}{N}}{\sigma\sqrt{\frac{T}{N}}}\Big)$$
$$\longrightarrow \Phi(0)=\frac{1}{2}. \tag{1.2}$$
Here Φ is the CDF of a standard normal distribution.

1.2.2. *Binomial Model.* In an N-period binomial model, the stock price process is specified by three parameters:
$$p,u,d.$$

From time n to time $n+1$, the price jumps from S_n to uS_n with probability p and to dS_n with probability $1-p$. Equivalently, $\ln\frac{S_{n+1}}{S_n}$ takes value $\ln u$ with probability p and takes value $\ln d$ with probability $1-p$. Put
$$X_{n+1}=\frac{\ln\frac{S_{n+1}}{S_n}-\frac{\ln u+\ln d}{2}}{\frac{\ln u-\ln d}{2}}.$$
Then
$$X_{n+1}=\begin{cases}1, & \text{with probability } p\\ -1, & \text{with probability } 1-p\end{cases}.$$
The stock return from time n to time $n+1$ is now given by
$$\ln\frac{S_{n+1}}{S_n}=\alpha+\lambda X_{n+1},$$
where, introduced for notational simplicity,
$$\alpha=\frac{\ln u+\ln d}{2},\qquad \lambda=\frac{\ln u-\ln d}{2}.$$
That is, from time n *to* $n+1$, *the return of stock is a scaled one-step random walk with drift.* Moreover, the stock

(1) expected rate of return

$$\mathbb{E}\Big[\ln\frac{S_{n+1}}{S_n}\Big] = \alpha + \lambda\big(p - (1-p)\big) = \alpha + \lambda(2p-1),$$

(2) variance of rate of return

$$\mathbb{V}\Big[\ln\frac{S_{n+1}}{S_n}\Big] = \lambda^2\big(1 - (2p-1)^2\big) = 4\lambda^2 p(1-p),$$

(3) probability to increase

$$\mathbb{P}(\text{stock price increases in this period}) = \mathbb{P}\Big(\ln\frac{S_{n+1}}{S_n} > 0\Big) = p.$$

1.3. **Approximate the BS model by binomial models.** Let's cut $[0, T]$ into N infinitesimal intervals with $N+1$ time slots:

$$\begin{array}{ccccccccccc} t=0 & \longrightarrow & t=\frac{T}{N} & \longrightarrow & t=\frac{2T}{N} & \longrightarrow & \dots & \longrightarrow & t=\frac{(N-1)T}{N} & \longrightarrow & t=T \\ n=0 & \longrightarrow & n=1 & \longrightarrow & n=2 & \longrightarrow & \dots & \longrightarrow & n=N-1 & \longrightarrow & n=N \end{array}$$

By the analysis in Section 1.2, using Binomial models to approximate the BS model means that we are using a scaled one-step random walk with drift to approximate a normal distribution *over each infinitesimal time interval*. After all, this means that over each infinitesimal time interval, we simplify a normal distribution by cutting it into two halves: the right of its center is signaled by stepping to the right in the scaled one-step random walk with drift, while the left of its center is signaled by stepping to the left.

In order to facilitate the approximation reasonably, we need to specify the three parameters for each N:[1]

$$p_N, u_N, d_N.$$

Comparing the formulas for the two cases in Section 1.2, we would require:

$$\frac{\mu T}{N} = \alpha_N + \lambda_N(2p_N - 1) + o\left(\frac{1}{N}\right) \tag{1.3}$$

$$\frac{\sigma^2 T}{N} = 4\lambda_N^2 p_N(1-p_N) + o\left(\frac{1}{N}\right) \tag{1.4}$$

$$P_N \to p \in (0,1), \tag{1.5}$$

[2] where

---

[1]We now put on the subindex N to indicate the reliance of the parameters on N.

[2]In (1.3), we want $\frac{\mu T}{N} \approx \alpha_N + \lambda_N(2p_N - 1)$; but to make the approximation meaningful, the error must be immaterial compared to these quantities themselves. This leads to (1.3). In (1.5), we could require the stronger condition $P_N \to \frac{1}{2}$ in view of (1.2), but the precise value of p won't matter as will be seen in the next section. An intuition comes from the fact that the physical probability will not affect security prices in binomial models. We rule out the case $P_N \to 0$ or 1 which means that over an infinitesimal amount of time the stock is nearly impossible to rise or to drop, which is sharply away from the BS model implication (1.2).

$$\alpha_N = \frac{\ln u_N + \ln d_N}{2}, \qquad \lambda_N = \frac{\ln u_N - \ln d_N}{2}. \tag{1.6}$$

Remark 1.1. In the above, we cared only at the infinitesimal time interval level. In fact, if the parameters are selected as in (1.3) and (1.4), the approximation at the terminal time is also good. Indeed, the terminal return of the stock over $[0, T]$ in the BS model is given by

$$\ln\left(\frac{S_T}{S_0}\right) = \mu T + \sigma B_T \quad \sim \quad \mathrm{N}\big(\mu T, \sigma^2 T\big).$$

The terminal return of the stock in the binomial model is given by

$$\ln \frac{S_N}{S_0} = \sum_{n=0}^{N-1} \ln \frac{S_{n+1}}{S_n} = \alpha_N N + \lambda_N \sum_{n=1}^{N} X_n.$$

Observe that X_n's are independent so that $\sum_{n=1}^N X_n$ is a random walk and is thus approximately normal (see Theorem 3.1 in Appendix at the end of this note)[3]. Thus $\ln \frac{S_N}{S_0}$ is approximately normal with mean $\alpha_N N + \lambda_N(2p_N - 1)N$, which is approximately μT by (1.3), and variance $4\lambda_N^2 p_N(1 - p_N)N$, which is approximately $\sigma^2 T$ by (1.4).

A final word about interest rate. A dollar grows to e^{rT} in the BS model and to $(1 + r_N)^N$ in the binomial model. So we select r_N such that:

$$(1 + r_N)^N = e^{rT}. \tag{1.7}$$

2. Limiting to the BS pricing formula

2.1. Consider an N-period binomial model with a stock and cash involving four parameters (p_N, u_N, d_N, r_N) *specified as in* (1.3)-(1.7). The price of a European call option on the stock with strike price K is given by

$$\begin{aligned} V_0^N =& \frac{1}{(1 + r_N)^N} \sum_{j=0}^{N} \binom{N}{j} \widetilde{p_N}^j (1 - \widetilde{p_N})^{N-j} \big(u_N^j d_N^{N-j} S_0 - K\big)^+ \\ =& e^{-rT} \sum_{j=0}^{N} \binom{N}{j} \widetilde{p_N}^j (1 - \widetilde{p_N})^{N-j} \big(u_N^j d_N^{N-j} S_0 - K\big)^+, \end{aligned}$$

where

$$\widetilde{p_N} = \frac{1 + r_N - d_N}{u_N - d_N}.$$

[3]The reader may figure out that $\sum_{n=1}^N \frac{X_n + 1}{2}$ is just an S_n in Theorem 3.1.

Note that

$$\begin{aligned}
&(u_N^j d_N^{N-j} S_0 - K)^+ > 0 \\
\iff & u_N^j d_N^{N-j} S_0 - K > 0 \\
\iff & j \ln u_N + (N-j)\ln d_N + \ln S_0 > \ln K \\
\iff & j > \frac{-N \ln d_N + \ln \frac{K}{S_0}}{\ln u_N - \ln d_N} := M_N.
\end{aligned}$$

Therefore,

$$\begin{aligned}
V_0^N =& e^{-rT} \sum_{j > M_N} \binom{N}{j} \widetilde{p_N}^j (1-\widetilde{p_N})^{N-j} (u_N^j d_N^{N-j} S_0 - K) \\
=& e^{-rT} S_0 \sum_{j > M_N} \binom{N}{j} \widetilde{p_N}^j (1-\widetilde{p_N})^{N-j} u_N^j d_N^{N-j} - e^{-rT} \sum_{j > M_N} \binom{N}{j} \widetilde{p_N}^j (1-\widetilde{p_N})^{N-j} K \\
:=& e^{-rT} S_0 \cdot U_1 - e^{-rT} K \cdot U_2.
\end{aligned}$$

We shall invoke the Central Limit Theorem 3.1 to compute the limit when $N \to \infty$. For example, for $\lim_{N\to\infty} U_2$, one apply the theorem (with $\le$ replaced by $>$) with $\widetilde{p_N}$ and

$$t_N := \frac{M_N - N\widetilde{p_N}}{\sqrt{N\widetilde{p_N}(1-\widetilde{p_N})}}.$$

2.2. Computing the limits. For computational convenience, we shall keep α_N and λ_N but skip the direct use of u_N and d_N, since we have direct control of α_N and λ_N via (1.3) and (1.4). Note first that

$$\ln u_N = \alpha_N + \lambda_N, \qquad \ln d_N = \alpha_N - \lambda_N. \tag{2.1}$$

Therefore,

$$\widetilde{p_N} = \frac{1 + r_N - d_N}{u_N - d_N} = \frac{e^{\frac{rT}{N}} - e^{\alpha_N - \lambda_N}}{e^{\alpha_N + \lambda_N} - e^{\alpha_N - \lambda_N}} = \frac{e^{\frac{rT}{N} - \alpha_N + \lambda_N} - 1}{e^{2\lambda_N} - 1}.$$

Solving (1.3) and (1.4), we have

$$\lambda_N = \sqrt{\frac{\frac{\sigma^2 T}{N} + o(\frac{1}{N})}{4p_N(1-p_N)}} = \frac{\sigma\sqrt{T}}{\sqrt{4p(1-p)}} \frac{1}{\sqrt{N}} \big(1 + o(1)\big) \longrightarrow 0, \tag{2.2}$$

$$\alpha_N = \frac{\mu T}{N} - \lambda_N(2p_N - 1) + o(\frac{1}{N}) \longrightarrow 0, \tag{2.3}$$

where the second equality in the first equation follows from (1.5). Therefore,

$$\begin{aligned}
&\lim_{N\to\infty} N\lambda_N(p_N - \widetilde{p_N}) \\
=&\lim_{N\to\infty} N\lambda_N \left(p_N - \frac{e^{\frac{rT}{N}-\alpha_N+\lambda_N}-1}{e^{2\lambda_N}-1}\right) = \lim_{N\to\infty} N\lambda_N \left(p_N - \frac{e^{\frac{(r-\mu)T}{N}+2\lambda_N p_N + o(\frac{1}{N})}-1}{e^{2\lambda_N}-1}\right) \\
=&\lim_{N\to\infty} N\lambda_N \frac{p_N(e^{2\lambda_N}-1) - e^{\frac{(r-\mu)T}{N}}\cdot e^{2\lambda_N p_N}\cdot e^{o(\frac{1}{N})}+1}{e^{2\lambda_N}-1} \\
=&\lim_{N\to\infty} N\lambda_N \frac{p_N\left(1+\frac{2\lambda_N}{1!}+\frac{4\lambda_N^2}{2!}+o(\frac{1}{N})-1\right) -}{2\lambda_N + o(\frac{1}{N})} \\
&\qquad \frac{-(1+\frac{(r-\mu)T}{N}+o(\frac{1}{N}))(1+\frac{2\lambda_N p_N}{1!}+\frac{4\lambda_N^2 p_N^2}{2!}+o(\frac{1}{N}))(1+o(\frac{1}{N}))+1}{con't} \\
=&\lim_{N\to\infty} N\lambda_N \frac{-\frac{(r-\mu)T}{N}+2\lambda_N^2 p_N(1-p_N)+o(\frac{1}{N})}{2\lambda_N+o(\frac{1}{N})} \\
=&\lim_{N\to\infty} N \frac{-\frac{(r-\mu)T}{N}+\frac{\sigma^2 T}{2N}+o(\frac{1}{N})}{2} \\
=&\frac{\sigma^2 T}{4} - \frac{(r-\mu)T}{2},
\end{aligned}$$

where in the second last equality we used (1.4). In particular, since $\lim_{N\to\infty} N\lambda_N = \infty$, it follows that

$$\lim_{N\to\infty}(p_N - \widetilde{p_N}) = 0,$$

so that

$$\lim_{N\to\infty}\widetilde{p_N} = p.$$

Consequently, the previous estimate of $N\lambda_N(p_N - \widetilde{p_N})$ together with (2.2) and (2.3) yields

$$\begin{aligned}
t_N =&\frac{M_N - N\widetilde{p_N}}{\sqrt{N\widetilde{p_N}(1-\widetilde{p_N})}} = \frac{\frac{-N\ln d_N + \ln\frac{K}{S_0}}{\ln u_N - \ln d_N} - N\widetilde{p_N}}{\sqrt{N\widetilde{p_N}(1-\widetilde{p_N})}} = \frac{\frac{-N(\alpha_N-\lambda_N)+\ln\frac{K}{S_0}}{2\lambda_N} - N\widetilde{p_N}}{\sqrt{N\widetilde{p_N}(1-\widetilde{p_N})}} \\
=&\frac{-N(\alpha_N-\lambda_N) - 2\lambda_N N\widetilde{p_N}}{2\lambda_N\sqrt{N\widetilde{p_N}(1-\widetilde{p_N})}} + \frac{\ln\frac{K}{S_0}}{2\lambda_N\sqrt{N\widetilde{p_N}(1-\widetilde{p_N})}} \\
=&\frac{-N\left(\frac{\mu T}{N} - \lambda_N 2p_N + o(\frac{1}{N})\right) - 2\lambda_N N\widetilde{p_N}}{2\lambda_N\sqrt{N\widetilde{p_N}(1-\widetilde{p_N})}} + \frac{\ln\frac{K}{S_0}}{2\lambda_N\sqrt{N\widetilde{p_N}(1-\widetilde{p_N})}} \\
=&\frac{-\mu T + 2\lambda_N N(p_N - \widetilde{p_N}) + o(1)}{2\big(\lambda_N\sqrt{N}\big)\sqrt{\widetilde{p_N}(1-\widetilde{p_N})}} + \frac{\ln\frac{K}{S_0}}{2\big(\lambda_N\sqrt{N}\big)\sqrt{\widetilde{p_N}(1-\widetilde{p_N})}} \\
\longrightarrow&\frac{-\mu T + 2(\frac{\sigma^2 T}{4} - \frac{(r-\mu)T}{2})}{2\frac{\sigma\sqrt{T}}{\sqrt{4p(1-p)}}\sqrt{p(1-p)}} + \frac{\ln\frac{K}{S_0}}{2\frac{\sigma\sqrt{T}}{\sqrt{4p(1-p)}}\sqrt{p(1-p)}} \\
=&-\frac{2r-\sigma^2}{2\sigma}\sqrt{T} + \frac{\ln\frac{K}{S_0}}{\sigma\sqrt{T}}
\end{aligned}$$

$$:= -d_2.$$

Therefore, by Theorem 3.1,

$$\lim_{N\to\infty} U_2 = \frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty} e^{-\frac{s^2}{2}}\,\mathrm{d}s = 1-\Phi(-d_2) = \Phi(d_2).$$

For U_1, we shall do a small trick and reduce to the previous case. Note first that

$$\begin{aligned}
U_1 &= \sum_{j>M_N}\binom{N}{j}\widetilde{p_N}^j(1-\widetilde{p_N})^{N-j}u_N^j d_N^{N-j}\\
&= \sum_{j>M_N}\binom{N}{j}(\widetilde{p_N}u_N)^j\Big((1-\widetilde{p_N})d_N\Big)^{N-j}\\
&= \sum_{j>M_N}\binom{N}{j}[\widetilde{p_N}u_N+(1-\widetilde{p_N})d_N]^N\Big(\frac{\widetilde{p_N}u_N}{\widetilde{p_N}u_N+(1-\widetilde{p_N})d_N}\Big)^j\Big(\frac{(1-\widetilde{p_N})d_N}{\widetilde{p_N}u_N+(1-\widetilde{p_N})d_N}\Big)^{N-j}\\
&=e^{rT}\sum_{j>M_N}\binom{N}{j}(\widetilde{P_N}^*)^j(1-\widetilde{P_N}^*)^{N-j},
\end{aligned}$$

where we used the fact

$$\widetilde{p_N}u_N+(1-\widetilde{p_N})d_N = 1+r_N$$

so that

$$[\widetilde{p_N}u_N+(1-\widetilde{p_N})d_N]^N = (1+r_N)^N = e^{rT},$$

and put

$$\widetilde{P_N}^* := \frac{\widetilde{p_N}u_N}{\widetilde{p_N}u_N+(1-\widetilde{p_N})d_N} = \widetilde{p_N}\frac{u_N}{1+r_N} = \widetilde{p_N}\frac{u_N}{e^{\frac{rT}{N}}} \longrightarrow p\cdot\frac{1}{1} = p.$$

Furthermore, using (2.2) and (2.3) again,

$$\begin{aligned}
&\sqrt{N}(\widetilde{P_N}-\widetilde{P_N}^*)\\
&=\sqrt{N}\widetilde{P_N}\Big(1-\frac{u_N}{e^{\frac{rT}{N}}}\Big) = \sqrt{N}\widetilde{P_N}\frac{e^{\frac{rT}{N}}-u_N}{e^{\frac{rT}{N}}} = \sqrt{N}\widetilde{P_N}\frac{e^{\frac{rT}{N}}-e^{\alpha_N+\lambda_N}}{e^{\frac{rT}{N}}}\\
&=\sqrt{N}\widetilde{P_N}\frac{e^{\frac{rT}{N}}-e^{\frac{\mu T}{N}+2\lambda_N(1-p_N)+o(\frac{1}{N})}}{e^{\frac{rT}{N}}}\\
&=\sqrt{N}\widetilde{P_N}\frac{e^{\frac{rT}{N}}-e^{\frac{\mu T}{N}}\cdot e^{2\lambda_N(1-p_N)}\cdot e^{O(\frac{1}{N})}}{e^{\frac{rT}{N}}}\\
&=\sqrt{N}\widetilde{P_N}\frac{\Big(1+O(\frac{1}{N})\Big)-\Big(1+O(\frac{1}{N})\Big)\cdot\Big(1+2\lambda_N(1-p_N)+O(\frac{1}{N})\Big)\cdot\Big(1+O(\frac{1}{N})\Big)}{e^{\frac{rT}{N}}}\\
&=\sqrt{N}\widetilde{P_N}\frac{-2\lambda_N(1-p_N)+O(\frac{1}{N})}{1+O(\frac{1}{N})}\\
\longrightarrow&-\frac{2p(1-p)\sigma\sqrt{T}}{\sqrt{4p(1-p)}}\\
&=-\sigma\sqrt{T}\sqrt{p(1-p)}.
\end{aligned}$$

Therefore,

$$\begin{aligned} t_N^* =& \frac{M_N - N\widetilde{p_N}^*}{\sqrt{N\widetilde{p_N}^*(1-\widetilde{p_N}^*)}} = \frac{M_N - N\widetilde{p_N}}{\sqrt{N\widetilde{p_N}^*(1-\widetilde{p_N}^*)}} + \frac{N(\widetilde{p_N} - \widetilde{p_N}^*)}{\sqrt{N\widetilde{p_N}^*(1-\widetilde{p_N}^*)}} \\ \longrightarrow & - d_2 - \sigma\sqrt{T} \\ := & - d_1. \end{aligned}$$

Finally, by Theorem 3.1 again, we have

$$\lim_{N\to\infty} U_1 = e^{rT}\frac{1}{\sqrt{2\pi}}\int_{-d_1}^{\infty} e^{-\frac{s^2}{2}}\mathrm{d}s = e^{rT}\big(1-\Phi(-d_1)\big) = e^{rT}\Phi(d_1).$$

2.3. Summary. Putting together $\lim_{N\to\infty} U_2$ and $\lim_{N\to\infty} U_1$ yields the celebrated Black-Scholes option pricing formula:

$$\lim_N V_0^N = S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2),$$

where

$$\begin{aligned} d_2 = & \quad \tfrac{2r-\sigma^2}{2\sigma}\sqrt{T} - \tfrac{\ln\frac{K}{S_0}}{\sigma\sqrt{T}} & = \tfrac{1}{\sigma\sqrt{T}}\Big(\ln\tfrac{S_0}{K} + rT\Big) - \tfrac{\sigma}{2}\sqrt{T} \\ d_1 = & \quad d_2 + \sigma\sqrt{T} & = \tfrac{1}{\sigma\sqrt{T}}\Big(\ln\tfrac{S_0}{K} + rT\Big) + \tfrac{\sigma}{2}\sqrt{T} \end{aligned}$$

2.4. Comments on selection of parameters. All our specifications of the parameters p_N, u_N and d_N are contained in (1.3)-(1.5), which leaves lots of room of parameter selection and shows that the binomial approximation is very robust. Given any specification about p_N, we get a specification of λ_N and α_N via (2.2)-(2.3) and thus a specification of u_N and d_N via (2.1). For example, one may take the classical choice of

$$p_N = \frac{1}{2} + \frac{\mu}{2\sigma}\sqrt{\frac{T}{N}}.$$

Taking the $o(\frac{1}{N})$ in (2.2) to be $-\mu^2\frac{T^2}{N^2}$, one has $\lambda_n = \sigma\sqrt{\frac{T}{N}}$. Taking the $o(\frac{1}{N})$ in (2.3) to be 0, one has $\alpha_N = 0$. Thus

$$u_N = \frac{1}{d_N} = e^{\sigma\sqrt{\frac{T}{N}}}.$$

Of course, one does not have to make P_N close to $\frac{1}{2}$, and can freely play with the small o's!

3. Appendix: An elementary Central Limit Theorem

3.1. We include a classical proof of the following version of the CLT without resorting to characteristic functions.

Theorem 3.1 (Variant of Laplace's CLT)**.** *Suppose* $(p_n) \subset (0,1)$ *converges to some* $p \in (0,1)$. *For each* $n \in \mathbb{N}$, *let* S_n *be a random variable following the binomial distribution* $\mathrm{Bi}(n, p_n)$. *Then if* $t_n \longrightarrow t \in \mathbb{R}$,

$$\mathbb{P}\left(\frac{S_n - np_n}{\sqrt{np_n(1-p_n)}} \le t_n\right) \longrightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{s^2}{2}} \mathrm{d}s = \Phi(t).$$

3.2. Before proceeding to the proof, we review a few facts.

Lemma 3.2 (Stirling's Formula)**.** *For any* $n \in \mathbb{N}$,

$$n! = \sqrt{2\pi n} n^n e^{-n} e^{\phi_n}, \quad \textit{where} \quad 0 < \phi_n < \frac{1}{12}.$$

The following simple estimates follow from Taylor expansions.

Lemma 3.3. (1) *For* $t \in (-\frac{1}{2}, \frac{1}{2})$,

$$\left|\frac{1}{\sqrt{1+t}} - 1\right| \le 5t.$$

(2) *For any* $x \in (-1, 1)$,

$$\left|\ln(1+x) - \left(x - \frac{x^2}{2}\right)\right| \le \frac{x^3}{3}$$

(3) *For any* $x \in (-1, 1)$,

$$\left|e^x - 1)\right| \le 2x.$$

3.3. Now fix any $a, b \in \mathbb{R}$ with $a < b$. In what follows, we use C to indicate quantities whose absolute values dominated by constants depending only on a, b and p. In different places, their precise values may differ, but they are immaterial for our estimates.

In all the arguments below, keep in mind that we may take n very large. Keep checking that all the estimates involving constants C hold for n large, and that the absolute values of all C's, regardless of n and j, have a common upper bound. For example, one may make n large such that $p_n > \frac{p}{2}$ and $q_n > \frac{q}{2}$. Here $q_n := 1 - p_n$ and $q := 1 - p$.

Claim 1: For $0 \le j \le n$, if $x_{nj} := \dfrac{j - np_n}{\sqrt{np_nq_n}} \in [a, b]$, then

$$\mathbb{P}(S_n = j) = \frac{1}{\sqrt{2\pi np_nq_n}} e^{-\frac{x_{nj}^2}{2}} \left(1 + \frac{C}{\sqrt{n}}\right).$$

Proof of Claim 1. By Stirling's Formula,

$$\begin{aligned}\mathbb{P}(S_n = j) &= \binom{n}{j} p_n^j (1-p_n)^{n-j} = \frac{n!}{j!(n-j)!} p_n^j (1-p_n)^{n-j} \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{j(n-j)}} \left(\frac{np_n}{j}\right)^j \left(\frac{nq_n}{n-j}\right)^{n-j} e^{\phi_n - \phi_j - \phi_{n-j}}. \qquad (3.1)\end{aligned}$$

Since $j = np_n + x_{nj}\sqrt{np_nq_n}$ and $n - j = nq_n - x_{nj}\sqrt{np_nq_n}$, we have

$$j > \frac{pn}{3}, \quad \text{and} \quad n - j > \frac{qn}{3} \quad \text{for } n \text{ large.}$$

Thus

$$\big|\phi_n - \phi_j - \phi_{n-j}\big| \le \frac{1}{12}\Big(\frac{1}{n} + \frac{1}{j} + \frac{1}{n-j}\Big) = \frac{C}{n},^{4}$$

and

$$e^{\phi_n-\phi_j-\phi_{n-j}} = 1 + \frac{C}{n}. \tag{3.2}$$

Moreover,

$$\begin{aligned}\frac{j(n-j)}{n} =& n\Big(p_n + x_{nj}\sqrt{\frac{p_nq_n}{n}}\Big)\Big(q_n - x_{nj}\sqrt{\frac{p_nq_n}{n}}\Big)\\ =& n\left(p_nq_n + (q_n-p_n)x_{nj}\sqrt{\frac{p_nq_n}{n}} - x_{nj}^2\frac{p_nq_n}{n}\right)\\ =& n\Big(p_nq_n + \frac{C}{\sqrt{n}}\Big) = np_nq_n + C\sqrt{n},\end{aligned}$$

[5] so that by Lemma 3.3(1),

$$\begin{aligned}\sqrt{\frac{n}{j(n-j)}} =& \frac{1}{\sqrt{np_nq_n}}\frac{1}{\sqrt{1+\frac{C}{\sqrt{n}p_nq_n}}} = \frac{1}{\sqrt{np_nq_n}}\Big(1 + \frac{C}{\sqrt{n}p_nq_n}\Big)\\ =& \frac{1}{\sqrt{np_nq_n}}\Big(1+\frac{C}{\sqrt{n}}\Big).\end{aligned} \tag{3.3}$$

We also have by Lemma 3.3(2),

$$\begin{aligned}&\ln\left(\frac{np_n}{j}\right)^j\left(\frac{nq_n}{n-j}\right)^{n-j}\\ =& -j\ln\frac{j}{np_n} - (n-j)\ln\frac{n-j}{nq_n}\\ =& -(np_n + x_{nj}\sqrt{np_nq_n})\ln\left(1 + \frac{x_{nj}\sqrt{q_n}}{\sqrt{np_n}}\right) - (nq_n - x_{nj}\sqrt{np_nq_n})\ln\left(1 - \frac{x_{nj}\sqrt{p_n}}{\sqrt{nq_n}}\right)\\ =& -(np_n + x_{nj}\sqrt{np_nq_n})\Big[\frac{x_{nj}\sqrt{q_n}}{\sqrt{np_n}} - \frac{x_{nj}^2q_n}{2np_n} + \frac{C}{n^{\frac{3}{2}}}\Big]\\ &\qquad - (nq_n - x_{nj}\sqrt{np_nq_n})\Big[\frac{x_{nj}\sqrt{p_n}}{\sqrt{nq_n}} - \frac{x_{nj}^2p_n}{2nq_n} + \frac{C}{n^{\frac{3}{2}}}\Big]\\ =& -\frac{x_{nj}^2}{2} + \frac{C}{\sqrt{n}},\end{aligned}$$

so that by Lemma 3.3(3)

$$\left(\frac{np_n}{j}\right)^j\left(\frac{nq_n}{n-j}\right)^{n-j} = e^{-\frac{x_{nj}^2}{2}}e^{\frac{C}{\sqrt{n}}} = e^{-\frac{x_{nj}^2}{2}}\Big(1+\frac{C}{\sqrt{n}}\Big). \tag{3.4}$$

---

[4]For example, $|C| \le \frac{1}{12}(1+\frac{3}{p}+\frac{3}{q})$ will work for large n.

[5]For example, $|C| \le |a| \vee |b| + a^2 \vee b^2$ will work for large n.

Inserting (3.2)-(3.4) into (3.1), we have

$$\begin{aligned}\mathbb{P}(S_n = j) =& \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{np_nq_n}} e^{-\frac{x_{nj}^2}{2}} \Big(1 + \frac{C}{\sqrt{n}}\Big)\Big(1 + \frac{C}{\sqrt{n}}\Big)\Big(1 + \frac{C}{n}\Big) \\ =& \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{np_nq_n}} e^{-\frac{x_{nj}^2}{2}} \Big(1 + \frac{C}{\sqrt{n}}\Big),\end{aligned}$$

here all the C's may be different and depend on n, j, but all their absolute values have a *common* dominant $M > 0$ depending *only* on a, b, p.

3.4. Claim 2:

$$\mathbb{P}\left(a \le \frac{S_n - np_n}{\sqrt{np_n(q_n}} \le b\right) \longrightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{s^2}{2}} \mathrm{d}s.$$

Proof of Claim 2.

$$\begin{aligned}&\mathbb{P}\left(a \le \frac{S_n - np_n}{\sqrt{np_nq_n}} \le b\right) \\ =& \sum_{j: a \le x_{nj} \le b} \mathbb{P}(S_n = j) \\ =& \sum_{j: a \le x_{nj} \le b} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{np_nq_n}} e^{-\frac{x_{nj}^2}{2}} \Big(1 + \frac{C}{\sqrt{n}}\Big).\end{aligned}$$

Note that since x_{nj}'s are points in $[a, b]$ differing by $\frac{1}{\sqrt{np_nq_n}}$ for adjacent j's, $\sum_{j:a\le x_{nj}\le b} \frac{1}{\sqrt{np_nq_n}} e^{-\frac{x_{nj}^2}{2}}$ is a Riemann sum for the integral $\int_a^b e^{-\frac{s^2}{2}} \mathrm{d}s$ with partition $\Gamma_n = \{x_{nj}\}$. Since $|\Gamma| = \frac{1}{\sqrt{np_nq_n}} \longrightarrow 0$, we have $\sum_{j:a\le x_{nj}\le b} \frac{1}{\sqrt{np_nq_n}} e^{-\frac{x_{nj}^2}{2}} \longrightarrow \int_a^b e^{-\frac{s^2}{2}} \mathrm{d}s$. And thus

$$\Big| \sum_{j: a \le x_{nj} \le b} \frac{1}{\sqrt{np_nq_n}} e^{-\frac{x_{nj}^2}{2}} \frac{C}{\sqrt{n}} \Big| \le \frac{M}{\sqrt{n}} \sum_{j: a \le x_{nj} \le b} \frac{1}{\sqrt{np_nq_n}} e^{-\frac{x_{nj}^2}{2}} \longrightarrow 0.$$

This proves Claim 2.

3.5. Claim 3. For any $t \in \mathbb{R}$,

$$\mathbb{P}\left(\frac{S_n - np_n}{\sqrt{np_n(1-p_n)}} \le t\right) \longrightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{s^2}{2}} \mathrm{d}s.$$

Proof of Claim 3. Fix $t \in \mathbb{R}$. Pick any $\varepsilon > 0$. Since $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{s^2}{2}} \mathrm{d}s = 1$, we can take $a_0, b_0 \in \mathbb{R}$ such that $a_0 < t < b_0$ and

$$\frac{1}{\sqrt{2\pi}} \int_{a_0}^{b_0} e^{-\frac{s^2}{2}} \mathrm{d}s > 1 - \varepsilon.$$

In particular,

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a_0} e^{-\frac{s^2}{2}} \mathrm{d}s = 1 - \frac{1}{\sqrt{2\pi}} \int_{a_0}^{\infty} e^{-\frac{s^2}{2}} \mathrm{d}s < \varepsilon.$$

For simplicity, write $X_n = \frac{S_n - np_n}{\sqrt{np_nq_n}}$. Since $\mathbb{P}(a_0 \le X_n \le b_0) \longrightarrow \frac{1}{\sqrt{2\pi}}\int_{a_0}^{b_0} e^{-\frac{s^2}{2}}\mathrm{d}s$ by Claim 2, there exists $N_1 \in \mathbb{N}$ such that

$$\mathbb{P}(a_0 \le X_n \le b_0) \ge \frac{1}{\sqrt{2\pi}}\int_{a_0}^{b_0} e^{-\frac{s^2}{2}}\mathrm{d}s - \varepsilon > 1 - 2\varepsilon \quad \text{for any } n \ge N_1,$$

so that

$$\mathbb{P}(X_n < a_0) = 1 - \mathbb{P}(X_n \ge a_0) \le 2\varepsilon \quad \text{for all } n \ge N_1.$$

Now since $\mathbb{P}(a_0 \le X_n \le t) \longrightarrow \frac{1}{\sqrt{2\pi}}\int_{a_0}^{t} e^{-\frac{s^2}{2}}\mathrm{d}s$, there exists $N_2 \in \mathbb{N}$ such that

$$\left|\mathbb{P}(a_0 \le X_n \le t) - \frac{1}{\sqrt{2\pi}}\int_{a_0}^{t} e^{-\frac{s^2}{2}}\mathrm{d}s\right| < \varepsilon \quad \text{for all } n \ge N_2.$$

Thus for any $n \ge N_1 \vee N_2$,

$$\begin{aligned}&\left|\mathbb{P}(X_n \le t) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{s^2}{2}}\mathrm{d}s\right| \\ =&\left|\mathbb{P}(X_n < a_0) + \mathbb{P}(a_0 \le X_n \le t) - \frac{1}{\sqrt{2\pi}}\int_{a_0}^{t} e^{-\frac{s^2}{2}}\mathrm{d}s - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a_0} e^{-\frac{s^2}{2}}\mathrm{d}s\right| \\ \le& 2\varepsilon + \varepsilon + \varepsilon.\end{aligned}$$

This proves Claim 3.

3.6. Proof of Theorem 3.1. Take $\delta > 0$ such that

$$\frac{1}{\sqrt{2\pi}}\int_{t-\delta}^{t+\delta} e^{-\frac{s^2}{2}}\mathrm{d}s < \varepsilon.$$

Take $N \in \mathbb{N}$ such that for any $n \ge N$,

$$\left|\mathbb{P}(X_n \le t+\delta) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t+\delta} e^{-\frac{s^2}{2}}\mathrm{d}s\right| < \varepsilon,$$

$$\left|\mathbb{P}(X_n \le t-\delta) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t-\delta} e^{-\frac{s^2}{2}}\mathrm{d}s\right| < \varepsilon.$$

Since $t_n \longrightarrow t$, there exists $N > N_1$ such that $t_n \in (t-\delta, t+\delta)$ for all $n \ge N$. Thus if $n \ge N$,

$$\mathbb{P}(X_n \le t_n) \le \mathbb{P}(X_n \le t+\delta) \le \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t+\delta} e^{-\frac{s^2}{2}}\mathrm{d}s + \varepsilon \le \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{s^2}{2}}\mathrm{d}s + 2\varepsilon,$$

$$\mathbb{P}(X_n \le t_n) \ge \mathbb{P}(X_n \le t-\delta) \ge \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t-\delta} e^{-\frac{s^2}{2}}\mathrm{d}s - \varepsilon \ge \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{s^2}{2}}\mathrm{d}s - 2\varepsilon,$$

so that

$$\left|\mathbb{P}(X_n \le t_n) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{s^2}{2}}\mathrm{d}s\right| \le 2\varepsilon.$$

We are done!