1. (a) The distribution will be approximately normal by the central limit theorem, since n is large and $np = 800$ is also large. The mean is $np = 800$ points and standard deviation is $2\sqrt{np(1-p)} = 12.65$ points.
(b) The probability is $P(P \geq 820) \approx P(N_{800,12.65} > 819.5) \approx 0.0616$.
(c) The distribution will be approximately normal by the central limit theorem, since the number of shots in each case is large. Mean is $1 \cdot 1000 \cdot 0.8 + 2 \cdot 2000 \cdot 0.4 + 3 \cdot 500 \cdot 0.2 = 2700$ points, variance is $1^2 \cdot 1000 \cdot 0.8 \cdot 0.2 + 2^2 \cdot 2000 \cdot 0.4 \cdot 0.6 + 3^2 \cdot 500 \cdot 0.2 \cdot 0.8 = 2800$ so $\sigma = \sqrt{2800} \approx 52.92$ points.
(d) The probability is $P(2500 \leq P \leq 2700) \approx P(2499.5 < N_{2700,52.92} < 2700.5) \approx 0.5037$.

---

2. (a) Average A is normal with mean $\mu_A = 10.05$g and $\sigma_A = 0.10/\sqrt{10} = 0.0316$g
Average B is normal with mean $\mu_B = 10.05g$ and $\sigma_B = 0.10/\sqrt{25} = 0.02$g.
(b) $P(N_{10.05,0.10} > 10.15) = 0.1586$ (c) $1 - 0.8414^{10} = 0.8223$ (d) $\mu = 100.5$g and $\sigma = 0.3162$g (e) $P(N_{100.5,0.316} > 101) = 0.0569$ (f) $P(N_{10.05,0.0316} < 9.95) = 0.00078$
(g) Difference is normal with mean $\mu_{A-B} = \mu_A - \mu_B = 0$ and $\sigma_{A-B} = \sqrt{\sigma_A^2 + \sigma_B^2} = 0.0374$g (h) $P(N_{0,0.0374} > 0.05) = 0.0907$.

---

3. (a) 2/7 (b) $\mu = 3.5$min, $\sigma = 2.0207$min (c) Approximately normal by central limit theorem with $\mu = 3.5$min and $\sigma = 0.2333$min (d) $P(N_{3.5,0.2333} > 3.57) = 0.3821$ (e) Number of times is binomial with $n = 75$, $p = 2/7$. Normal approximation has $\mu = np = 21.43$, $\sigma = \sqrt{np(1-p)} = 3.91$ yielding estimate $P(N_{21.43,3.91} > 29.5) = 0.0196$. Exact answer is $P(B_{75,2/7} \geq 30) = 0.0220$.

---

4. (a) Distribution is Poisson (it is counting rarely-occurring events) with parameter $\lambda = 3$, $p_X(n) = e^{-\lambda}\lambda^n/n!$
(b) $P(P_3 = 0) = e^{-3} \approx 0.0498$ (c) $P(P_3 \geq 5) \approx 0.1847$ (d) Number of interruptions today is Poisson with $\lambda = 3/7$ so probability is $P(P_{3/7} > 0) = 1 - e^{-3/7} \approx 0.3486$ (e) Number is Poisson with $\lambda = 90/7$ so probability is $P(P_{90/7} = 15) = e^{-90/7}(90/7)^{15}/15! \approx 0.0865$

---

5. (a) Distribution is exponential (it is a memoryless wait time) with $\lambda = 1/30$, so $p_Y(y) = \lambda e^{-\lambda y}$ for $y \geq 0$
(b) $P(E_{1/30} > 30) = e^{-1} \approx 0.3689$ (c) $P(E_{1/30} < 10) = 1 - e^{-1/3} \approx 0.2835$

---

6. (a) Binomial with $n = 130$, $p = 0.02$, so $0.98^{130} + \binom{130}{1}0.02^1 0.98^{129} = 0.2643$
(b) Parameter is the average number who like the course, which is $130 \cdot 0.02 = 2.6$. The resulting probability estimate is $P(X = 0) + P(X = 1) = e^{-2.6} + 2.6e^{-2.6} \approx 0.2674$.
(c) Parameters are the mean and standard deviation, which are $np = 2.6$ and $\sqrt{np(1-p)} = 1.5962$. The continuity-corrected probability estimate is $P(-0.5 \leq N \leq 1.5) \approx 0.2193$.
(d) The Poisson is better since the normal approximation to the binomial is not especially good for np small. On the other hand, this is precisely the situation where the Poisson distribution is a good approximation.

---

7. (a) (i) $\theta^4(1-\theta)^{14}$ (ii) $\frac{4}{\theta} - \frac{14}{1-\theta}$ (iii) $\hat{\theta} = 4/18$.
(b) (i) $\theta^{14}e^{-3\theta}/518400$ (ii) $\frac{14}{\theta} - 3$ (iii) $\hat{\theta} = 14/3$.
(c) (i) $\theta^{-3}e^{-8/\theta}$ (ii) $\frac{-3}{\theta} + \frac{8}{\theta^2}$ (iii) $\hat{\theta} = 8/3$.
(d) (i) $e^{-(1-\theta)^2-(-2-\theta)^2-(0-\theta)^2}\pi^{-3/2}$ (ii) $2(1-\theta) + 2(-2-\theta) + 2(0-\theta)$ (iii) $\hat{\theta} = -1/3$.
(e) (i) $4(\theta-2)(\theta-5)/\theta^4$ (ii) $\frac{1}{\theta-2} + \frac{1}{\theta-5} - \frac{4}{\theta}$ (iii) $\hat{\theta} = 8$ (there is a second root $\hat{\theta} = 5/2$ but it is less than 5).

---

8. (a) $E(\hat{\lambda}_1) = \frac{1}{2}[E(x_1) + E(x_2)] = \lambda$, unbiased, $\text{var}(\hat{\lambda}_1) = \frac{1}{4}[\text{var}(x_1) + \text{var}(x_2)] = \lambda/2$.

   (b) $E(\hat{\lambda}_2) = \frac{1}{4}[E(x_1) + E(x_2) + E(x_3) + E(x_4)] = \lambda$, unbiased
   $\text{var}(\hat{\lambda}_2) = \frac{1}{16}[\text{var}(x_1) + \text{var}(x_2) + \text{var}(x_3) + \text{var}(x_4)] = \lambda/4$.

   (c) $E(\hat{\lambda}_3) = E(x_1) - E(x_2) + 2E(x_3) = 2\lambda$, biased
   $\text{var}(\hat{\lambda}_3) = \text{var}(x_1) + \text{var}(x_2) + 4\text{var}(x_3) = 6\lambda$.

   (d) $E(\hat{\lambda}_4) = \frac{1}{5}[E(x_1) + E(x_2) + 2E(x_3) + E(x_4)] = \lambda$, unbiased,
   $\text{var}(\hat{\lambda}_4) = \frac{1}{25}[\text{var}(x_1) + \text{var}(x_2) + 4\text{var}(x_3) + \text{var}(x_4)] = 7\lambda/25$.

   (e) The estimator $\hat{\lambda}_2$ is the most efficient since its variance is the smallest.

---

9. (a) By expected value properties, $E(\hat{\mu}) = E(ax) + E(by) = aE(x) + bE(y) = (a + 3b)\mu$,
   By variance properties, $\text{var}(\hat{\mu}) = \text{var}(ax) + \text{var}(by) = a^2\text{var}(x) + b^2\text{var}(y) = 4a^2 + 3b^2$.

   (b) Need $(a + 3b)\mu = \mu$ for unbiased, so $a = 1 - 3b$. Then $\text{var}(\hat{\mu}) = 4(1 - 3b)^2 + 3b^2 = 39b^2 - 24b + 4$ has derivative $78b - 24$ which is zero for $b = 4/13$, which is a min. Therefore $a = 1/13$, $b = 4/13$ gives the minimum.

---

10. (a) H_0: $\mu_c = \mu_p$, H_a: $\mu_c > \mu_p$, $\sigma = \sqrt{\frac{1}{330} + \frac{1}{200}} = 0.0896$, p-value is $P(N_{0,0.0896} > -0.02) = 0.5883$, fail to reject H_0.
   Alternatively, with H_0: $\mu_c = \mu_p$, H_a: $\mu_c < \mu_p$, the p-value is $P(N_{0,0.0896} < -0.02) = 0.4117$, fail to reject H_0.

   (b) H_0: $\mu_t = \mu_p$, H_a: $\mu_t > \mu_p$, $\sigma = \sqrt{\frac{1}{350} + \frac{1}{200}} = 0.0886$, p-value is $P(N_{0,0.0886} > 0.07) = 0.2147$, fail to reject H_0.

   (c) H_0 :$\mu_c = \mu_t$, H_a : $\mu_c \neq \mu_t$, $\sigma = \sqrt{\frac{1}{350} + \frac{1}{330}} = 0.0767$, p-value is $2P(N_{0,0.0767} > 0.09) = 0.2406$, fail to reject H_0.

   (d) The probability of a type I error in each case is 4%, so the probability of making at least one error is $1 - 0.96^3 \approx 11.5\%$, assuming independence. (A reasonable estimate is also $4\% + 4\% + 4\% = 12\%$.)

   (e) Publishing only statistically-significant outcomes creates a bias in the literature, especially since performing multiple tests increases the probability of a false-positive result, as calculated in (d). All tests performed should always be reported, and corrections for multiple tests must be included.

---

11. (a) H_0: $p_c = p_p$, H_a: $p_c > p_p$, $p_{\text{pool}} = 0.3134$, $\sigma_{\text{pool}} = 0.08015$, $\hat{p}_c - \hat{p}_p = 0.02985$, p-value is $P(N_{0,0.08015} > 0.02985) = 0.3548$, fail to reject H_0.
   95% CI for placebo: $(0.1889, 0.4081)$, 95% CI for crystal: $(0.2159, 0.4408)$. Note $\hat{p}_{\text{placebo}} = 0.2985$, $\sigma_{\text{placebo,prop}} = 0.0559$, $\hat{p}_{\text{crystal}} = 0.3284$, $\sigma_{\text{crystal,prop}} = 0.0574$.

   (b) H_0: $\mu_m = \mu_f$, H_a: $\mu_m > \mu_f$, $\mu_{m-f} = 17078$, $\sigma_{m-f} = 6708$, p-value is $P(N_{0,6708} > 17078) = 0.0054$, reject H_0.
   95% CI for male: $(102711, 114471)$, 95% CI for female: $(79753, 103273)$. Note $\sigma_{\text{male,avg}} = 30000/\sqrt{100} = 3000$ and $\sigma_{\text{female,avg}} = 30000/\sqrt{25} = 6000$.

   (c) H_0: $p_w = p_b$, H_a: $p_w > p_b$, $p_{\text{pool}} = 0.0808$, $\sigma_{\text{pool}} = 0.00783$, p-value is $P(N_{0,0.00783} > 0.03176) = 2.5 \cdot 10^{-5}$, reject H_0.
   95% CI for White: $(8.47\%, 10.82\%)$, 95% CI for Black: $(5.49\%, 7.45\%)$. Note $\hat{p}_{\text{White}} = 9.69\%$, $\sigma_{\text{White,prop}} = 0.60\%$, $\hat{p}_{\text{Black}} = 6.47\%$, $\sigma_{\text{Black,prop}} = 0.50\%$.

---

12. (a) H_0: $p = 0.156$, H_a: $p \neq 0.156$, $np = 24.8$, $\sqrt{np(1-p)} = 4.575$, p-value is $\approx 2P(N_{24.8,4.575} > 29.5) = 0.3043$, fail to reject H_0.
(b) H_0: $p = 0.156$, H_a: $p < 0.156$, $np = 24.8$, $\sqrt{np(1-p)} = 4.575$, p-value is $\approx P(N_{24.8,4.575} < 30.5) = 0.8934$, fail to reject H_0.
(c) H_0: $p = 0.156$, H_a: $p > 0.156$, $np = 24.8$, $\sqrt{np(1-p)} = 4.575$, p-value is $\approx P(N_{24.8,4.575} > 29.5) = 0.1521$, reject H_0 / fail to reject H_0 / fail to reject H_0.
(d) (i) One study cannot prove anything definitively. It also provides essentially zero evidence toward the hypothesis that hydroxychloroquine is effective in lowering the hospitalization rate, from test (b).
(ii) One study cannot prove anything definitively. It provides somewhat weak evidence toward the hypothesis that hydroxychloroquine increases the hospitalization rate, from test (c).
(iii) This is also not entirely accurate, because test (c) does provide weak evidence suggesting that hydroxychloroquine increases the hospitalization rate.
(e) The power of the test, and hence the strength of the conclusions, increases with a larger sample size. The p-values here would shift to 0.0752, 0.9624, 0.0376.

---

13. (a) $\mu = 0.38, \sigma = 0.0686$, 80%: $(0.2920, 0.4680)$, 90%: $(0.2671, 0.4929)$, 95%: $(0.2455, 0.5145)$, 99%: $(0.2032, 0.5568)$.
(b) 9 times as many students: 450 in total.
(c) Need $n = \frac{\hat{p}(1-\hat{p})}{(0.02/1.9600)^2} \approx 2262.7$ (so 2263). If $\hat{p}$ is unknown then worst case is $\hat{p} = 0.5$ with $n = 2401$.
(d) H_0: $p = 0.3$, H_a: $p \neq 0.3$, p-value is $P(|B_{50,0.3} - 15| \geq |19 - 15|) \approx 2P(N_{15,3.2404} > 18.5) = 0.2801$, fail to reject H_0.
(e) H_0: $p = 0.5$, H_a: $p < 0.5$, p-value is $P(B_{50,0.5} < 19) \approx P(N_{25,3.5355} < 19.5) = 0.0599$, reject / fail to reject / fail to reject H_0.
(f) (d-i) correct / correct / correct, (d-ii) type II / type II / type II,
(e-i) type I / correct / correct, (e-ii) correct / type II / type II.
(g) For $n = 50$, sample proportion $\hat{p}$ is approximately normal with $\mu = 0.46$ and $\sigma = \sqrt{0.46 \cdot 0.54/50} = 0.0705$.
Null hypothesis proposes p is approximately normal with $\mu = 0.5$ and $\sigma = \sqrt{0.5 \cdot 0.5/50} = 0.0707$.
At 11%, reject when $P(N_{0.5,0.0707} < x) = 0.11$ yielding $x = 0.4133$, so error probability is $\beta = P(N_{0.46,0.0705} > 0.4133) = 0.7463$ with power $1 - \beta = 0.2537$.
For $n = 500$, sample proportion $\hat{p}$ is approximately normal with $\mu = 0.46$ and $\sigma = \sqrt{0.46 \cdot 0.54/500} = 0.0223$.
Null hypothesis proposes p is approximately normal with $\mu = 0.5$ and $\sigma = \sqrt{0.5 \cdot 0.5/500} = 0.0224$.
At 11%, reject when $P(N_{0.5,0.0224} < x) = 0.11$ yielding $x = 0.4726$, so error probability is $\beta = P(N_{0.46,0.0223} > 0.4726) = 0.2863$ with power $1 - \beta = 0.7137$.

---

14. (a) $\mu = 1.6, \sigma = 4$, 80%: $(-0.6926, 3.8926)$, 90%: $-1.3425, 4.5425)$, 95%: $(-1.9062, 5.1062)$, 99.5: $(-3.4213, 6.6213)$.
(b) p-value is $2P(N_{0,4/\sqrt{5}} > 8/5) = 0.3711$, fail to reject at 10%, 3%, 1%.
(c) Sample mean $\overline{x}$ is normal with $\mu = 4$ and $\sigma = 4/\sqrt{5}$.
At 10%, reject when $P(N_{0,4/\sqrt{5}} > x) = 0.10$ yielding $x = 2.2925$, so error probability is $\beta = P(N_{4,4/\sqrt{5}} < 2.2925) = 0.1699$ with power $1 - \beta = 0.8301$.
At 3%, reject when $P(N_{0,4/\sqrt{5}} > x) = 0.03$ yielding $x = 3.3645$, so error probability is $\beta = P(N_{4,4/\sqrt{5}} < 3.3645) = 0.3612$ with power $1 - \beta = 0.6388$.
At 1%, reject when $P(N_{0,4/\sqrt{5}} > x) = 0.01$ yielding $x = 4.1615$, so error probability is $\beta = P(N_{4,4/\sqrt{5}} < 4.1615) = 0.5360$ with power $1 - \beta = 0.4640$.
(d) Sample mean $\overline{x}$ is normal with $\mu = 4$ and $\sigma = 4/\sqrt{20}$.
At 10%, reject when $P(N_{0,4/\sqrt{20}} > x) = 0.10$ yielding $x = 1.1463$, so error probability is $\beta = P(N_{4,4/\sqrt{20}} < 1.1463) = 0.0007$ with power $1 - \beta = 0.9993$.
At 3%, reject when $P(N_{0,4/\sqrt{20}} > x) = 0.03$ yielding $x = 1.6822$, so error probability is $\beta = P(N_{4,4/\sqrt{20}} < 1.6822) = 0.0048$ with power $1 - \beta = 0.9952$.
At 1%, reject when $P(N_{0,4/\sqrt{20}} > x) = 0.01$ yielding $x = 2.0808$, so error probability is $\beta = P(N_{4,4/\sqrt{20}} < 2.0808) = 0.0159$ with power $1 - \beta = 0.9840$.
