18.05 Problem Set 5, Spring 2025 Solutions

Problem 1. (20 pts: 5, 5, 5, 5 pts.) **Random walks**

(a) Solution: Let $Y_i \sim$ Bernoulli(0.5) so that $Y = Y_1 + \cdots + Y_n \sim$ binomial$(n, 0.5)$. Then $X_i = 2Y_i - 1$, and $S_n = X_1 + \cdots + X_n = 2Y - n$ is the total displacement, with distribution symmetric about zero. Hence,

$$P(|S_n| > 100) = 2 \cdot P(S_n > 100) = 2 \cdot P(2Y - n > 100) = 2 \cdot P(Y > 50 + n/2)$$

Letting $n = 10000$, this probability is $2(1 - \texttt{pbinom(5050, 10000, 0.5)}) = \boxed{0.3124.}$

(b) Solution: We have $E[X_i] = 0$ and $\text{Var}(X_i) = 1$. So $E[S_{10,000}] = 0$, and because each step is independent, $\text{Var}(S_{10,000}) = 10000$. The CLT says $S_{10000} \sim N(0, 100^2)$. Hence,

$$P(|S_{10000}| > 100) = P\left(\frac{S_{10000}}{100} > 1\right) \approx P(|Z| \geq 1) = 2(1 - P(Z \leq 1)) \approx \boxed{0.3173.}$$

The estimate is very good!

(c) Solution: Now we are given the probability and need to find the n. Since $\text{Var}(S_n) = n$, as in (b) we have:

$$0.9 = P(|S_{10000}| > 100) = P\left(\frac{S_{10000}}{\sqrt{n}} > \frac{100}{\sqrt{n}}\right) \approx P(|Z| \geq 100/\sqrt{n}) = 2(1 - P(Z \leq 100/\sqrt{n}))$$

Solving, we need $P(Z \leq 100/\sqrt{n}) = 0.55$, which implies $100/\sqrt{n} = \texttt{qnorm(0.55)} = 0.12566$. So $n = (100/0.12566)^2 \approx \boxed{633300}$ seconds. So just over a week!

Xeno almost surely crossed the 100 step barrier for the first time much sooner. The problem is that he keeps wandering back toward where he started, cancelling out his progress.

(d) Solution: Now we have:

X_i values:	1	0	-1
probabilities:	1/3	1/3	1/3

So $E[X_i] = 0$ and $\text{Var}(X) = E[X^2] - E[X]^2 = \frac{(-1)^2}{3} + \frac{0^2}{3} + \frac{1^2}{3} - 0^2 = \frac{2}{3}$. The standard deviation is $\sqrt{2/3}$.

We know that $E[S_{10,000}] = 0$, and because each step is independent, $\text{Var}(S_{10,000}) = 10,000 \cdot \frac{2}{3}$. The standard deviation is about 81.650.

The Central Limit Theorem tells us that $S_{10000} \sim N(0, (81.65)^2)$. Hence,

$$P\left(\frac{S_{10000}}{81.65} > \frac{100}{81.65}\right) \approx 1 - P(|Z| \leq 1.22) = 2(1 - P(Z \leq 1.22)) \approx \boxed{0.221.}$$

The probability of a large displacement went down. This makes sense because the variance of each step went down due to the chance of resting.

Problem 2. (20 pts: 5,5,5,5 pts.) **Fat tails**

(a) Solution: (i) The mean of a uniform distribution is its midpoint, so the range should be centered at 39.9, i.e. $[39.9 - c, 39.9 + c]$.

The uniform(0,1) distribution has width 1 and standard deviation $1/\sqrt{12}$. Our uniform distribution has width $2c$, hence standard deviation $2c/\sqrt{12}$. (Since scaling a random variable scales its standard deviation). Thus

$$\frac{2c}{\sqrt{12}} = 12.0 \;\Rightarrow\; c = 12.0\sqrt{3} \approx 20.785$$

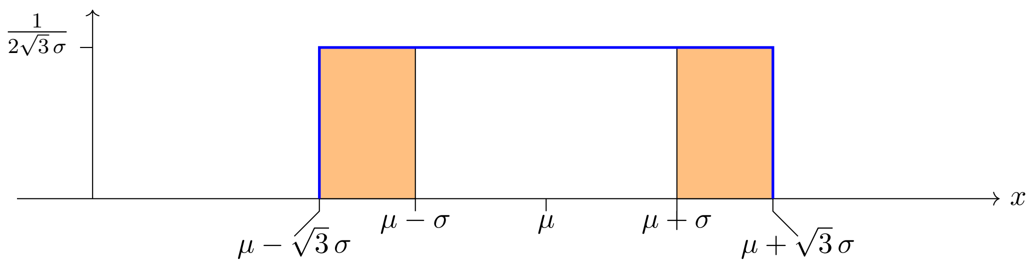
Thus, the range is $39.9 \pm 12.0\sqrt{3}$ or approximately $[19.115, 60.685]$.

(ii) From part (i) we know that the full range of possible values of $x - \mu$ is

$$-12.0\sqrt{3} \le x - \mu \le 12.0\sqrt{3}.$$

In terms of σ, this is $-\sigma\sqrt{3} \le x - \mu \le \sigma\sqrt{3}$.

A 1σ event means $|x - \mu| \ge \sigma$. Here's a graph with the 1σ event shaded. To emphasize that the picture is the same for any uniform random variable, we use labels μ and σ instead of decimals.



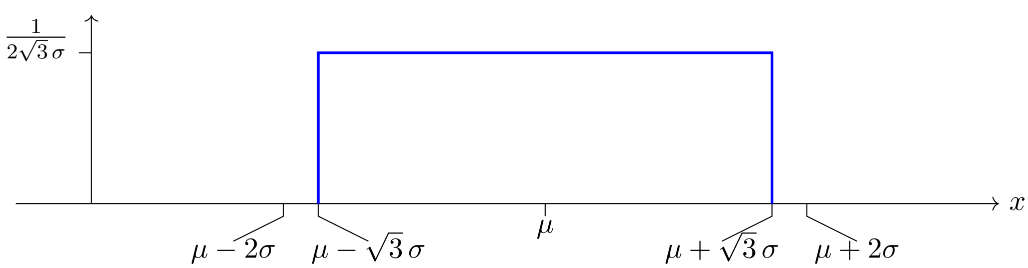
The 1σ event is shaded.

The width of the range of uniform random variable is $2\sqrt{3}\,\sigma$. So the constant density has value $f(x) = \dfrac{1}{2\sqrt{3}\,\sigma}$.

Thus, the probability of the 1σ event, which is the shaded area, is

$$\text{Prob. of } 1\sigma \text{ event } = \frac{1}{2\sqrt{3}\,\sigma}\left(2(\sqrt{3}-1)\sigma\right) = \boxed{1 - 1/\sqrt{3} \approx 0.423.}$$

(iii) Since the range of the uniform is $\mu \pm \sqrt{3}\,\sigma$, $\mu \pm 2\sigma$ is outside the range. That is, there are no 2σ events. So the probability is 0.



$\mu \pm 2\sigma$ is outside the range of the distribution.

(iv) The top of the range of possible temperatures is $\mu + \sqrt{3}\sigma = 39.9 + \sqrt{3} \cdot 12.0 = 58.85$. Since 76 is above this range, this event has zero probability.

(b)

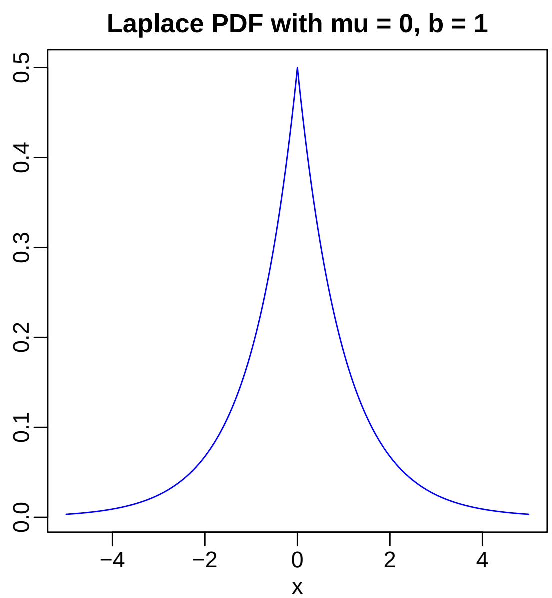
We use the `pnorm` function in R to compute this. Here is the code and results.

```
mu = 39.9
sig = 12.0
ans_5aii = pnorm(mu - sig, mu, sig) + 1 - pnorm(mu + sig, mu, sig)
ans_5aiii = pnorm(mu - 2*sig, mu, sig) + 1 - pnorm(mu + 2*sig, mu, sig)
ans_5aiv = 1 - pnorm(76, mu, sig)
```

Results: (ii) 0.317 (iii) 0.0455 (iv) 0.00131, note that part (iv) is only looking for the extreme on one side.

Note: Since the problem is expressed in terms of standard deviations from the mean, we could have used the standard normal distribution.
That is: `ans_6aii = pnorm(-1) + 1 - pnorm(1)`.
Or, using the symmetry of the normal distribution: `ans_6bii = 2*pnorm(-1)`.

(c) Solution: (o) Here is the plot done in R.



Laplace density with $\mu = 0$, $b = 1$.

(i) Looking on Wikipedia, we see mean = μ and variance = $2b^2$. So $\sigma = \sqrt{2}\,b$. Thus, we choose $\boxed{\mu = 39.9}$ and $\boxed{b = 12.0/\sqrt{2} \approx 8.485.}$ The cdf is

$$F(x) = \begin{cases} \frac{1}{2}\mathrm{e}^{(x-\mu)/b} & \text{for } x \le \mu \\ 1 - \frac{1}{2}\mathrm{e}^{-(x-\mu)/b} & \text{for } x \ge \mu \end{cases}$$

The integral for the mean is

$$\text{mean } = \frac{1}{2b}\int_{-\infty}^{\infty} x\mathrm{e}^{-|x-\mu|/b}\,dx.$$

Once we've found the mean is μ (what else could it be?) the integral for variance is

$$\text{var } = \frac{1}{2b}\int_{-\infty}^{\infty} (x-\mu)^2\mathrm{e}^{-|x-\mu|/b}\,dx.$$

The integral for $F(x)$ is

$$F(x) = \frac{1}{2b}\int_{-\infty}^{x} \mathrm{e}^{-|u-\mu|/b}\, du.$$

For mean and variance, the change of variable $y = (x - \mu)/b$ will simplify the integrals. They can then be computed using integration by parts. In both the mean and variance, you can exploit symmetry to make the calculation even easier.

For the cdf, the cases $x < \mu$ and $x > \mu$ are handled separately. Again the change of variable $y = (u - \mu)/b$ will make things simpler.

(ii-iv) We use the formula for $F(x)$ and the values of μ and b found in part (i)

$$\begin{aligned} P(1\sigma \text{ event}) &= F(\mu - \sigma) + 1 - F(\mu + \sigma) \approx 0.243 \\ P(2\sigma \text{ event}) &= F(\mu - 2\sigma) + 1 - F(\mu + 2\sigma) \approx 0.0591 \\ P(76) = 1 - F(76) &= \frac{1}{2}\mathrm{e}^{-(76-\mu)/b} \approx 0.00710 \end{aligned}$$

Note: We could have used symmetry to get, e.g. $P(1\sigma \text{ event}) = 2F(\mu - \sigma)$.

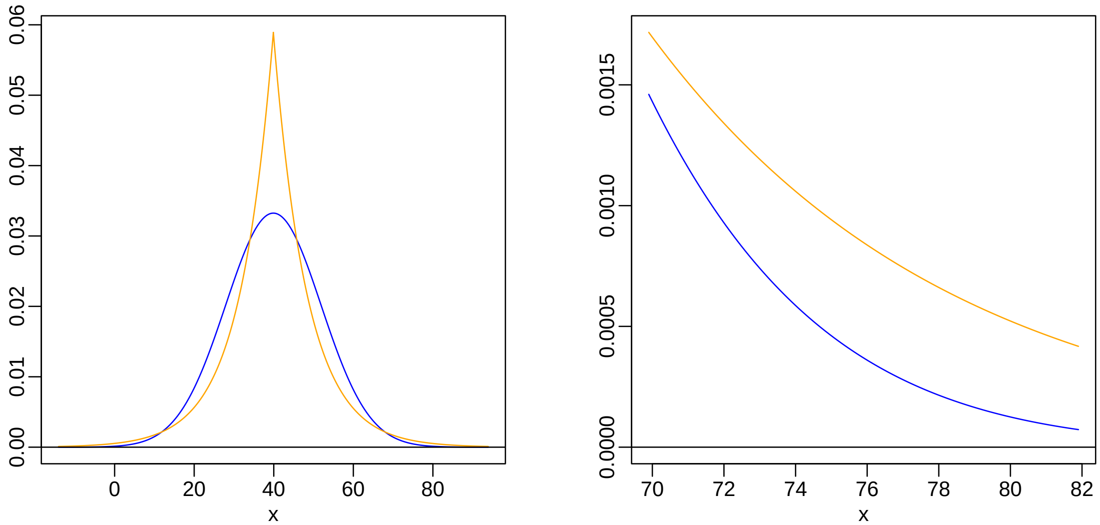
(d) **Solution:** Note: A temperature above 76 degrees is at least 3σ above the mean high temperature. That is, it is in the 'right tail' of the distribution. Thus, we want to decide which distribution is most likely to have a value in its tail.

According to part (a), for the uniform distribution a 3σ deviation is impossible (has zero probability). Thus, this distribution provides the worst explanation.

According to part (b), for the normal distribution this event has an approximate probability of 0.00131. That is, on average, there will be a 76 degree in February about once every 760 years.

According to part (c), for the double-exponential distribution this event has an approximate probability of 0.00710. That is, on average, there will be a 76 degree in February about once every 140 years.

Of the two, the Laplace probability seems closer to what would really happen. That is, because of its much thicker tails, the Laplace transformation provides a better explanation of extreme data – which we suspect occurs more often than the tiny probability given by the normal distribution. Another factor here is distributional shift due to climate change.



Laplace and normal both with $\mu = 39.9$, $\sigma = 12.0$.

The figure above on the right shows the right tails of the Laplace and normal distributions. The Laplace is the higher one (in orange).

The moral: If you see only a moderately extreme event (say, 2σ), the normal distribution and the double-exponential distribution can be comparably good explanations. For example, their respective probabilities of a 2σ event are similar, at approximately 0.046 and 0.059. However, if you see an extreme event, be skeptical that the underlying distribution is the normal distribution, despite its convenience in other ways. Its tails are just so thin.

Problem 3. (15: 5, 5, 5 pts.) **Maximum likelihood estimates**

(a) Solution: The hypotheses are that the urn used is urn 1, urn 2 or urn 3. The data is that the chosen balls were red, then green, then red. Call the data RGR. So,

$$P(RGR|\text{urn }1) = \frac{5\cdot 5\cdot 4}{12\cdot 11\cdot 10} \approx 0.0758$$
$$P(RGR|\text{urn }2) = \frac{3\cdot 8\cdot 2}{15\cdot 14\cdot 13} \approx 0.0176$$
$$P(RGR|\text{urn }3) = \frac{7\cdot 7\cdot 6}{17\cdot 16\cdot 15} \approx 0.0721$$

So, the maximum likelihood estimate is that the urn is urn 1.

(b) Solution: The likelihood for x_i is $f(x_i|\lambda) = \dfrac{\lambda^{10}}{9!} x_i^9 \mathrm{e}^{-\lambda x_i}$. So, the likelihood of the data is

$$f(\text{data}|\lambda) = \prod_{i=1}^{m} f(x_i|\lambda) = \frac{\lambda^{10m}}{(9!)^m} P^9 \mathrm{e}^{-\lambda S},$$

where $P = \prod x_i$ (product of data) and $S = \sum x_i$ (sum of data).

So, the log likelihood is $l(\lambda) = 10m\ln(\lambda) + 9\ln(P) - \lambda S - m\ln(9!)$. Taking the derivative and setting it to 0, we get

$$l'(\lambda) = \frac{10m}{\lambda} - S = 0 \;\Rightarrow\; \boxed{\hat{\lambda} = \frac{10m}{S}.}$$

Note: The distribution mean is $10/\lambda$ and $\hat{\lambda} = 10/(S/m) = 10/\overline{x}$, where $\overline{x}$ is the data mean.

(c) Solution: The pdf of a uniform$(0, b)$ distribution takes two values

$$f(x|b) = \begin{cases} 1/b & \text{if } x \text{ is in } [0, b] \\ 0 & \text{otherwise} \end{cases}$$

Since the likelihood is a product of the likelihoods of each data point the likelihood function is

$$f(2.5,\ 19.75,\ 12.0,\ 7.0|b) = \begin{cases} (1/b)^4 & \text{if all 4 data points are in the interval } [0, b] \\ 0 & \text{otherwise} \end{cases}$$

This is maximized when b is as small as possible while making sure all the data points are in $[0, b]$. This means b is the the maximum of the data, i.e. $\boxed{b = 19.75}$.

Problem 4. (20 pts: 5,5,5,5 pts.) **Monty Hall: Sober and drunk**

(a) Solution: In all three parts to this problem we have 3 hypotheses:

H_A = 'the car is behind door A'
H_B = 'the car is behind door B'
H_C = 'the car is behind door C'.

In all three parts the data is D = 'Monty opens door B and reveals a goat'.

The key to our Bayesian update table is the likelihoods: For part (a), since Monty is sober he always reveals a goat.

$P(D|H_A)$: H_A says the car is behind A. So, assuming H_A is true, Monty is equally likely to pick B or C and reveal a goat. Thus $P(D|H_A) = 1/2$.

$P(D|H_B)$: If H_B is true, the car is behind B and sober Monty will never choose B (and if he did it would not reveal a goat). So $P(D|H_B) = 0$.

$P(D|H_C)$: If H_c is true, the car is behind C. Since sober Monty doesn't make mistakes he will open door B and reveal a goat. So $P(D|H_C) = 1$.

Here is the table for this situation.

H	$P(H)$	$P(D|H)$	Bayes numer.	Posterior
H_A	1/3	1/2	1/6	1/3
H_B	1/3	0	0	0
H_C	1/3	1	1/3	2/3
Total:	1	–	1/2	1

Therefore, Aviva should switch to door C, since the probability H_C is true is twice that of H_A.

(b) Solution: Some of the likelihoods change in this setting.

$P(D|H_A)$: If H_A is true then the car is behind A. So Monty is equally likely to show B or C and reveal a goat. Thus $P(D|H_A) = 1/2$. (Remember, D is the 'Monty opens door B and reveals a goat'.)

$P(D|H_B)$: If H_B is true then the car is behind B, drunk Monty might show B, but if he does we won't reveal a goat. (He will ruin the game.) So $P(D|H_B) = 0$.

$P(D|H_C)$: H_c says the car is behind C. Drunk Monty is equally likely to show B or C. If he chooses B he'll reveal a goat. So $P(D|H_C) = 1/2$.

Our table is now:

H	$P(H)$	$P(D\|H)$	Bayes numer.	Posterior
H_A	1/3	1/2	1/6	1/2
H_B	1/3	0	0	0
H_C	1/3	1/2	1/6	1/2
Total:	1	–	1/3	1

So in this case switching is just as good (or as bad) as staying with the original choice.

(c) Solution: We have to recompute the likelihoods. Remember the data is that Monty chooses door B and reveals a goat.

$P(D|H_A)$: If the car is behind A then sober or drunk Monty is equally likely to choose door B and reveal a goat. Thus $P(D|H_A) = 1/2$.

$P(D|H_B)$: If the car is behind door B then if he chooses it he will reveal a car, not a goat. So the probability of the data given H_B is 0, i.e., $P(D|H_B) = 0$.

$P(D|H_C)$: Let S be the event that Monty is sober and S^c the event he is drunk. From the table in (a), we see that $P(D|H_C, S) = 1$ and from the table in (b), we see that $P(D|H_C, S^c) = 1/2$. Thus, by the law of total probability

$$P(D|H_C) = P(D|H_C, S)P(S) + P(D|H_C, S^c)P(S^c) = 1 \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{3}{4} = \frac{5}{8}$$

H	$P(H)$	$P(D\|H)$	Bayes numer.	Posterior
H_A	1/3	1/2	1/6	$4/9 = 0.444$
H_B	1/3	0	0	0
H_C	1/3	5/8	5/24	$5/9 = 0.556$
Total:	1	–	9/24	1

Thus, switching gives a probability of winning of approximately 0.556. So switching is still the best strategy. The intuitive feel for this is that even a little bit sober, Monty is giving some information by picking B, or, more precisely, avoiding C.

(d) Solution: We accept two answers: when $p = 0$ and 'never' (due to ambiguity of 'best'). Namely, the posterior will be 0.5 in the case when $p = 0$ for both switching and staying (cf. the table in (c)).

Problem 5. (20 pts: 5, 5, 5, 5 pts.) **Bayesian dice**

(a) Solution: We could solve this in one table by multiplying the likelihoods from each roll. But let's write out four tables to show the progression of the posterior.

hypothesis	prior	likelihood	Bayes numerator	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}=1\|\mathcal{H})$	$P(\mathcal{D}=1\|\mathcal{H})P(\mathcal{H})$	$P(\mathcal{H}\|\mathcal{D}=1)$
$\mathcal{H}_4$	1/5	1/4	1/20	7/20
$\mathcal{H}_6$	1/5	1/6	1/24	7/24
$\mathcal{H}_8$	1/5	1/8	1/32	7/32
$\mathcal{H}_{12}$	1/5	1/12	1/60	7/60
$\mathcal{H}_{20}$	1/5	1/20	1/100	7/100
total	1		1/7	1

After one roll, the 4-sided die is in the lead. We now move the posterior from the first table to the prior of the second table, and similarly for the third and fourth tables.

hypothesis	prior	likelihood	Bayes numerator	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}=7\|\mathcal{H})$	$P(\mathcal{D}=7\|\mathcal{H})P(\mathcal{H})$	$P(\mathcal{H}\|\mathcal{D}=7)$
$\mathcal{H}_4$	7/20	0	0	0
$\mathcal{H}_6$	7/24	0	0	0
$\mathcal{H}_8$	7/32	1/8	7/256	0.67406
$\mathcal{H}_{12}$	7/60	1/12	7/720	0.23966
$\mathcal{H}_{20}$	7/100	1/20	7/2000	0.086279
total	1		0.040565972	1

After two rolls, the 8-sided die is in the lead, because a roll of 7 ruled out the 4- and 6-sided dice. We may as well drop rows corresponding to hypotheses with zero probability since no future data can resuscitate them.

hypothesis	prior	likelihood	Bayes numerator	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}=1\|\mathcal{H})$	$P(\mathcal{D}=1\|\mathcal{H})P(\mathcal{H})$	$P(\mathcal{H}\|\mathcal{D}=1)$
$\mathcal{H}_8$	0.67406	1/8	0.084257	0.77625
$\mathcal{H}_{12}$	0.23966	1/12	0.019972	0.18400
$\mathcal{H}_{20}$	0.086279	1/20	0.0043140	0.039744
total	1		0.10854	1

After three rolls, the 8-sided die is even more in the lead.

hypothesis	prior	likelihood	Bayes numerator	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}=11\|\mathcal{H})$	$P(\mathcal{D}=11\|\mathcal{H})P(\mathcal{H})$	$P(\mathcal{H}\|\mathcal{D}=11)$
$\mathcal{H}_8$	0.77625	0	0	0
$\mathcal{H}_{12}$	0.18400	1/12	0.015333	0.88527
$\mathcal{H}_{20}$	0.039744	1/20	0.0019872	0.11473
total	1		0.017320	1

The fourth roll ruled out the 8-sided die, so in the end, the 12-sided die is most probable.

(b) **Solution:** Here we combine all four likelihoods into one table:

hypothesis	prior	likelihood	Bayes numerator	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}\|\mathcal{H})$	$P(\mathcal{D}\|\mathcal{H})P(\mathcal{H})$	$P(\mathcal{H}\|\mathcal{D})$
$\mathcal{H}_4$	1/5	$1/4 \cdot 3/4 \cdot 1/4 \cdot 3/4$	$3^2/4^4$	0.47189
$\mathcal{H}_6$	1/5	$1/6 \cdot 5/6 \cdot 1/6 \cdot 5/6$	$5^2/6^4$	0.25892
$\mathcal{H}_8$	1/5	$1/8 \cdot 7/8 \cdot 1/8 \cdot 7/8$	$7^2/8^4$	0.16058
$\mathcal{H}_{12}$	1/5	$1/12 \cdot 11/12 \cdot 1/12 \cdot 11/12$	$11^2/12^4$	0.078325
$\mathcal{H}_{20}$	1/5	$1/20 \cdot 19/20 \cdot 1/20 \cdot 19/20$	$19^2/20^4$	0.030285
total	1		0.074501	1

Now the MLE is the 4-sided die, which makes sense because all we know is that two of four rolls were 1s. We can no longer rule out the 4-, 6-, or 8-sided dice because the rolls of 7 and 11 are censored.

(c) Solution: The likelihood for n-sided is:

$$L(n) = \left(\frac{1}{n}\right)^5 \left(\frac{n-1}{n}\right)^{95}.$$

Since the prior is uniform, the posterior probabilities are then $L(n)/C$ where

$$C = L(4) + L(6) + L(8) + L(12) + L(20).$$

We arrive at the same answer with more effort using the full Bayes' table:

hypothesis	prior	likelihood	Bayes numerator	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}\|\mathcal{H})$	$P(\mathcal{D}\|\mathcal{H})P(\mathcal{H})$	$P(\mathcal{H}\|\mathcal{D})$
$\mathcal{H}_4$	1/5	$(1/4)^5(3/4)^{95}$	$(1/4)^5(3/4)^{95}/5$	3.7468×10^{-7}
$\mathcal{H}_6$	1/5	$(1/6)^5(5/6)^{95}$	$(1/6)^5(5/6)^{95}/5$	0.0010969
$\mathcal{H}_8$	1/5	$(1/8)^5(7/8)^{95}$	$(1/8)^5(7/8)^{95}/5$	0.026820
$\mathcal{H}_{12}$	1/5	$(1/12)^5(11/12)^{95}$	$(1/12)^5(11/12)^{95}/5$	0.29330
$\mathcal{H}_{20}$	1/5	$(1/20)^5(19/20)^{95}$	$(1/20)^5(19/20)^{95}/5$	0.67878
total	1		7.0452×10^{-10}	1

The MLE is the 20-sided die. This makes sense as 5 is the expected number of 1s in 100 rolls of a 20-sided die, whereas we would expect to see far more 1s with the other dice.

(d) Solution: Using the new prior, we have the Bayes' table:

hypothesis	prior	likelihood	Bayes numerator	posterior
$\mathcal{H}$	$P(\mathcal{H})$	$P(\mathcal{D}\|\mathcal{H})$	$P(\mathcal{D}\|\mathcal{H})P(\mathcal{H})$	$P(\mathcal{H}\|\mathcal{D})$
$\mathcal{H}_4$	96/100	$(1/4)^5(3/4)^{95}$	$(1/4)^5(3/4)^{95} \cdot 96/100$	3.59677×10^{-5}
$\mathcal{H}_6$	1/100	$(1/6)^5(5/6)^{95}$	$(1/6)^5(5/6)^{95}/100$	0.0010968
$\mathcal{H}_8$	1/100	$(1/8)^5(7/8)^{95}$	$(1/8)^5(7/8)^{95}/100$	0.026819
$\mathcal{H}_{12}$	1/100	$(1/12)^5(11/12)^{95}$	$(1/12)^5(11/12)^{95}/100$	0.29329
$\mathcal{H}_{20}$	1/100	$(1/20)^5(19/20)^{95}$	$(1/20)^5(19/20)^{95}/100$	0.67876
total	1		3.52273×10^{-11}	1

The MLE is still the 20-sided die. This makes sense because, while the new prior increased the posterior on the 4-sided die by a factor of nearly 100, the posterior is still extremely small due to the improbability of rolling only five 1s in 100 rolls of a 4-sided die.