1. I had a recent flight from Washington DC to Columbia. My plane had 66 seats on it and each seat was occupied with one passenger only. I asked the stewardess for a breakdown of the passengers by gender. She told me there were 36 females and 30 males.

(a) Suppose I took a sample of $n = 5$ passengers at random (and without replacement). How many samples of size 5 are possible?
(b) Let Y denote the number of females in a sample. What is the (named) distribution of Y? Is Y discrete or continuous? Explain.
(c) Calculate the probability that there are exactly 4 females in the sample. Note that if there 4 females in the sample, then the other passenger is a male.

2. An electrical product consists of n integrated circuits. Each circuit functions independently of the others. The probability that each circuit functions is p. Let r_n denote the reliability of the product.

- If the circuits are laid out in a **series**, the reliability of the product is
$$r_n = p^n.$$
- If the circuits are laid out in **parallel**, the reliability of the product is
$$r_n = 1 - (1-p)^n.$$

(a) Explain the difference between a series and parallel structure. Also, use simple probability arguments to explain why each of the formulas above is correct.
(b) If six circuits are laid out in parallel, how unreliable can the individual circuits be and still have the product's reliability to be at least $r_6 = 0.9999$?

3. The pH of a water sample taken from a specific lake is a continuous random variable Y with probability density function

$$f_Y(y) = \begin{cases} \frac{3}{8}(7-y)^2, & 5 \leq y \leq 7 \\ 0, & \text{otherwise.} \end{cases}$$

(a) I calculated the median of this distribution to be 5.413. Write out an equation (involving an integral) that when solved will give this answer (you don't have to solve the equation). Also, interpret (in words) what the median represents.
(b) Suppose I measure water samples independently (e.g., from different parts of the lake). I am interested in finding the first water sample with pH ≥ 6.0. What is the probability I will find such a water sample on the first or second sample tested?

4. A discrete random variable Y has the following probability mass function (pmf):

y	0	1	2	3	4
$p_Y(y)$	0.24	0.41	0.26	0.08	0.01

(a) Graph the cumulative distribution function (cdf). Neatness counts. Label axes.
(b) Suppose that $p_Y(y)$ is a probability model for Y, the number of times a student misses or is late to class per week (for a class that meets four times per week). The professor assigns weekly attendance scores W using the formula

$$W = 6 - 0.5Y - 0.25Y^2.$$

Find the mean weekly attendance score $E(W)$.

5. The time to failure (T, measured in hours) of a bearing used in a mechanical shaft is under investigation. The following times were recorded on $n = 84$ bearings:

27135.9	3293.5	16380.9	3406.6	12455.7	1014.4	4627.3
22800.0	425.5	30859.3	202.4	3508.5	5475.3	6924.4
534.9	288.6	1363.4	9200.5	154.3	325.2	4197.5
5592.2	682.1	172.3	25287.2	125.0	1753.7	1010.5
5118.9	2217.6	175.8	3142.7	5094.3	33154.8	427.5
6142.1	15180.4	971.4	103.2	2691.4	2406.8	3961.8
1814.8	1074.0	812.8	21022.6	4548.1	9877.8	707.9
1903.9	2293.7	2581.4	31597.1	25994.2	3661.3	2004.0
238.5	2164.7	22304.2	15.9	157.8	17673.4	143.0
672.2	2671.0	417.8	5421.3	290.9	17286.1	4263.6
8492.6	8885.1	16947.8	29890.5	4102.9	11009.1	15.3
5663.3	41.4	2002.8	1329.6	29821.9	139.0	130.6

One engineer assumes that T follows a Weibull distribution. Using the times above, I calculated the maximum likelihood estimates of β and η under the Weibull model to be

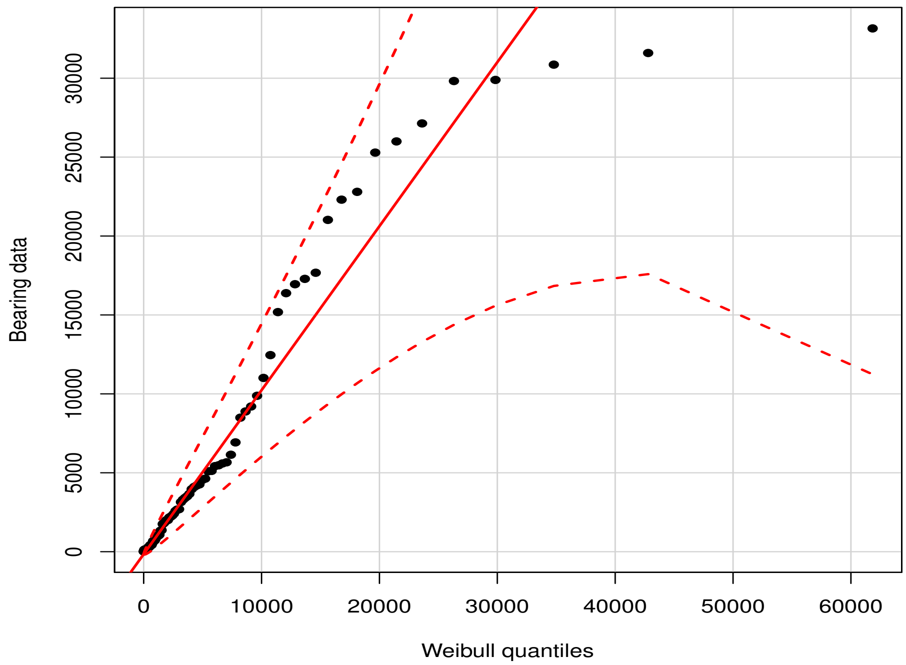
$$\begin{aligned}\widehat{\beta} &\approx 0.66 \\ \widehat{\eta} &\approx 5115.83\end{aligned}$$

Explain what is meant by the term "maximum likelihood estimates."
(b) Using the Weibull model with values $\widehat{\beta}$ and $\widehat{\eta}$ above, estimate the probability that a new bearing will fail before 15000 hours.
(c) Recall that the hazard function of T is

$$h_T(t) = \frac{f_T(t)}{S_T(t)},$$

where $S_T(t)$ is the survivor function of T. Calculate the estimated hazard function in this application. Also, what does the shape of the estimated hazard function suggest about the population of bearings?

(d) Here is a quantile-quantile plot for the bearing data (under a Weibull assumption):



Interpret what information is revealed by this plot.

Binomial:

$$p_Y(y) = \begin{cases} \binom{n}{y} p^y (1-p)^{n-y}, & y = 0, 1, 2, ..., n \\ 0, & \text{otherwise.} \end{cases}$$

Geometric:

$$p_Y(y) = \begin{cases} (1-p)^{y-1} p, & y = 1, 2, 3, ... \\ 0, & \text{otherwise.} \end{cases}$$

Negative binomial:

$$p_Y(y) = \begin{cases} \binom{y-1}{r-1} p^r (1-p)^{y-r}, & y = r, r+1, r+2, ... \\ 0, & \text{otherwise.} \end{cases}$$

Hypergeometric:

$$p_Y(y) = \begin{cases} \dfrac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}, & y \leq r \text{ and } n - y \leq N - r \\ 0, & \text{otherwise.} \end{cases}$$

Poisson:

$$p_Y(y) = \begin{cases} \dfrac{\lambda^y e^{-\lambda}}{y!}, & y = 0, 1, 2, ... \\ 0, & \text{otherwise.} \end{cases}$$

Exponential:

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y}, & y > 0 \\ 0, & \text{otherwise.} \end{cases} \qquad F_Y(y) = \begin{cases} 1 - e^{-\lambda y}, & y > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Gamma:

$$f_Y(y) = \begin{cases} \dfrac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y}, & y > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Normal (Gaussian):

$$f_Y(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma} e^{-(y-\mu)^2/2\sigma^2}, & -\infty < y < \infty \\ 0, & \text{otherwise.} \end{cases}$$

Weibull:

$$f_T(t) = \begin{cases} \dfrac{\beta}{\eta}\left(\dfrac{t}{\eta}\right)^{\beta-1} e^{-(t/\eta)^\beta}, & t > 0 \\ 0, & \text{otherwise.} \end{cases} \qquad F_T(t) = \begin{cases} 1 - e^{-(t/\eta)^\beta}, & t > 0 \\ 0, & \text{otherwise.} \end{cases}$$