Exam 2 Review with Solutions

STAT 381, Spr15

Instructor: Brian Powers

1 Memorize

If $X \sim Bern(p)$, $E(X) = p, Var(X) = pq$.
If $X \sim Binom(n,p)$, $f(x) = \binom{n}{x} p^x q^{n-x}$, $E(X) = np$, $Var(X) = npq$.
If $X \sim Poisson(\lambda t)$, $E(X) = Var(X) = \lambda t$.
If X_is are independent, $E(\bar{X}) = \mu, Var(\bar{X}) = \sigma^2/n$ regardless of the distribution.
CLT says: $\bar{X} \sim N(\mu, \sigma^2/n)$ approximately, or $(\bar{X} - \mu)/(\sigma/\sqrt{n}) \sim N(0,1)$
If $X \sim Unif_C(a,b)$, $f(x) = \frac{1}{b-a}$ for $x \in [a,b]$, $E(X) = \frac{a+b}{2}, Var(X) = \frac{(b-a)^2}{12}$.

2 Don't Have to Memorize

Poisson pmf: $f(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$ for $x \in \mathbb{N}$.
Hypergeometric, Geometric, Negative binomial pmf.
Normal distribution pdf
Gamma distribution pdf
t-distribution pdf

3 Review Problems

1. Give the name of the distribution best used to model each random variable:

   (a) Whether or not the first life savers in the package is pineapple
   Bernoulli

   (b) The number of cherry starburst candies in a package of 12
   Binomial

   (c) The number of mystery dum-dums you need to unwrap until you get a root beer
   Geometric

   (d) The number of trick-or-treaters until you've seen 10 Elsas (from Frozen)
   Negative Binomial

   (e) The number of red candies in a handful of 10 drawn from a bowl of red, blue and yellow candies
   Hypergeometric

   (f) The number of trick-or-treaters coming to your door between 8:05 and 8:10pm
   Poisson

   (g) The length of time between trick-or-treaters
   Exponential

(h) The length of time until 10 trick-or-treaters in total have arrived
Gamma

(i) The height of the next trick-or-treater
Normal

(j) The angle between the seconds hand and minutes hand when the first trick-or-treater arrives.
Uniform

2. In a game, you are allowed to roll a pair of fair dice 24 times. If you get a double six (i.e. you roll (6,6)) at least 4 times you win. What is the probability you win this game?
The probability of rolling two sixes is

$$P(66) = P(6)P(6) = \frac{1}{36}$$

Say X is the number of double-sixes rolled out of 24 trials. Then $X \sim Binom(24, \frac{1}{36})$. You win if $X \geq 4$, so

$$P(Win) = P(X \geq 4) = 1 - P(X \leq 3) = 1 - \texttt{binomcdf}(24, 1/36, 3) = .0041$$

3. Let X be the number of customer logins on a website during an hour. Assume X has a Poisson distribution with a mean of 120 login requests per hour.

(a) What is the probability that no one requests to log on the site in the next ten minutes?
10 minutes is $\frac{1}{6}$ of an hour, so the number of logins during 10 minutes will follow a Poisson with parameter $120/6 = 20$. Let $Y \sim Poisson(20)$.

$$P(Y = 0) = \texttt{poissonpdf}(20, 0) \approx 0$$

or

$$P(Y = 0) = \frac{20^0 e^{-20}}{0!} = e^{-20} \approx 0.$$

(b) Let W be the time in minutes between the 2nd and 3rd request. What is the distribution name of W? What is the expected value of W?
This is an exponential distribution (or Gamma). Since there is on average 2 login requests per minute, $\beta = 1/2$. The expected value is β, which is 30 seconds.

(c) Let T be the time in minutes until the 4th request. What is the distribution name of T? What is the expected value of T?
This is a Gamma distribution with $\alpha = 4, \beta = 1/2$. The expected value is $\alpha\beta = 2$ minutes.

4. At a certain gas station, the number of lottery winners each month follows a Poisson distribution with mean 1.

(a) What is the probability that there are no more than 3 winners this year?
The number of winners in a year will follow a Poisson distribution with mean 12. Let Y be the number of winners in a year.

$$P(Y \leq 3) = \texttt{poissoncdf}(12, 3) = .00229$$

(b) What is the probability there is at least one winner during a month?
Let X be the number of winners in a month.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \texttt{poissonpdf}(1, 0) = .63212$$

or

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{1^0 e^{-1}}{0!} = 1 - e^-1 = .63212$$

(c) What is the probability that there is at least one winner every month during the year?
Assuming whether it happens in each month is indepedent,

$$P(\geq 1 \text{ winner 12 months straight}) = (.63212)^{12} = .00407$$

5. Let $X_1, X_2, \ldots, X_{36}$ be a random sample from a continuous Uniform distribution over $[0, 2]$.

(a) Find the pdf for the population.
$f(x) = \frac{1}{2-0} = \frac{1}{2}$ for $x \in [0, 2]$, $f(x) = 0$ otherwise.

(b) Find the mean and variance for the population.
By theorem: $E(X) = \frac{b-a}{2} = \frac{2-0}{2} = 1$, $Var(X) = \frac{(b-a)^2}{12} = \frac{(2-0)^2}{12} = \frac{1}{3}$
You may also argue that $\mu = 1$ by integral,

$$\mu = \int_0^2 \frac{1}{2} dx = \frac{x}{2}\Big|_0^2 = 1 - 0 = 0$$

Or by symmetry of the density function about $x = 1$. You can calculate the variance directly by $Var(X) = E(X^2) - \mu^2$,

$$\sigma^2 = \int_0^2 \frac{x^2}{2} dx - 1^2 = \frac{x^3}{6}\Big|_0^2 - 1 = \frac{8}{6} - 1 = \frac{1}{3}$$

(c) Find the mean and variance for the sample mean $\bar{X} = (X_1 + \cdots + X_{36})/36$.
$E(\bar{X}) = \mu = 1$. $Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{1/3}{36} = \frac{1}{108}$

(d) Use the Central Limit Theorem to estimate $P(0.9 < \bar{X} < 1.1)$.
By Central Limit Theorem, $\bar{X} \sim N(1, \frac{1}{108})$ approximately. So

$$P(0.9 < \bar{X} < 1.1) \approx \texttt{normalcdf}(.9, 1.1, 1, \sqrt{1/108}) = .7013$$

or you can calculate $z_1 = (.9-1)/\sqrt{1/108} = -1.039$, $z_2 = (1.1-1)/\sqrt{1/108} = 1.039$,

$$P(-1.039 < Z < 1.039) = \texttt{normalcdf}(-1.039, 1.039) = .7012$$

6. If a r.v. X follows a normal distribution with mean 80 and standard deviation 20, find the following using the Empirical rule (68 - 95 - 99.7 rule).

   (a) $P(X \geq 20)$
   20 is 3 standard deviations from the mean. In the tail past 3 standard deviations there is approximately $.003/2 = .0015\%$. So $P(X \geq 20) = 1 - .0015 = .9985$

   (b) $P(|X - 80| \leq 40)$
   $P(|X - 80| \leq 40)$ is the probability of being within 2 standard deviations. That is .95 according to the empirical rule.

   (c) $P(X \leq 100)$
   100 is 1 standard deviation from the mean, there is $(1 - .68)/2 = .16$ probability in the upper tail, so $P(X \leq 100) = 1 - .16 = .84$.

   (d) $P(X > 120)$
   120 is 2 standard deviations above the mean, there is $.05/2 = .025$ probability in the upper tail, so $P(X > 120) = .025$.

7. The size of a download from a server follows a normal distribution with mean 4 MB and standard deviation 0.8 MB.

   (a) What is the probability that a download exceeds 5 MB?
   Let X be the size of a random download. $P(X > 5) = \texttt{normalcdf}(5, 1000, 4, .8) = .1056$

   (b) What are the 25th and 75th percentiles?
   The 25th percentile is $\texttt{invNorm}(.25, 4, .8) = 3.46$ MB. The 75th percentile is $\texttt{invNorm}(.75, 4, .8) = 4.54$ MB

   (c) If someone downloads 5 files, what is the probability that the total download is more than 22 MB?
   The sample sum will follow a Normal distribution with mean $5(4) = 20$ and standard deviation $\sqrt{5}(.8) = 1.789$. Let S_5 be the sample sum, with a sample size of 5.

$$P(S_5 > 22) = \texttt{normalcdf}(22, 10000, 20, 1.789) = .1318$$

8. Statistics show that on an average weekend night 1 out of 10 drivers on the road is drunk.

   (a) If 20 drivers are checked, what is the expected number of drunk drivers?
   The number of drunk drivers, X follows a binomial distribution with $n = 20, p = .1$, so $E(X) = np = 2$.

(b) If 20 drivers are checked, what is the probability that at least 1 of them is drunk?

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{20}{0}(.1)^0(.9)^{20} = .8784$$

or

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \texttt{binompdf}(20, .1, 0) = .8784$$

(c) If a sample of 200 drivers are checked, what is the mean and variance for the number of drunk drivers?
If $n = 200$, $\mu = 200(.1) = 20$, $\sigma^2 = 200(.1)(.9) = 18$